TECHNISCHE UNIVERSITÄT DARMSTADT

Kompetenzmodelle

MODELING AND TRAINING

A NEW CONCEPT OF

TEACHERS' DIAGNOSTIC COMPETENCE

D i s s e r t a t i o n

zur Erlangung der Doktorwürde (Dr. rer. nat.)

am Fachbereich Humanwissenschaften

der Technischen Universität Darmstadt

vorgelegt von

Dipl. Psych. Julia Klug,

geboren in Groß-Umstadt

Erstreferent: Prof. Dr. Bernhard Schmitz (Technische Universität Darmstadt)

Ko-Referent: Prof. Dr. Augustin Kelava (Technische Universität Darmstadt)

Tag der Einreichung: 25.08.2011

Tag der Disputation: 09.12.2011

Darmstadt 2011

D17

Für meine Eltern und meinen Opa Willibald

Acknowledgment

First of all, I would like to thank my supervisor Prof. Dr. Bernhard Schmitz for his support, guidance and faith in me since I worked as a student assistant and throughout the whole period of my doctoral thesis. He gave me the possibility to conduct my studies, let me present the results at national and international conferences and fostered me from the beginning leading to finally proposing me the job advertisement where I will now continue my academic career.

I also like to thank Prof. Dr. Augustin Kelava, who agreed to be the second reviewer of my thesis. He not only gave me valuable tips on modeling, but also supported me with a lot of humor to go the distance.

Furthermore, I am very thankful to Dr. Udo Keil. He was the one who first mentored me and stimulated my interest in research and methods. Udo, I guess, it is because of you that I want to be a researcher.

Very special thanks go to Dr. Simone Bruder with whom I worked together very closely. She supported me a lot and she was not only the best colleague ever, but she also is the best friend you can imagine.

Most importantly, I want to thank my parents, my grandfather Willibald and the rest of my family who always had faith in me and supported me with lots of love. Thank you!

I also like to thank my colleagues and friends who made work as enjoyable as possible. Particular thanks go to Florian Kattner, Anne-Kathrin Scheibe, Sylvana Keller, Sabine Ogrin, Mara Gerich, Monika Trittel, Kirsten van de Loo, Alice Ihringer, Henrik Bellhäuser, Oliver Christ and Jana Lampert.

Additionally, I would like to thank my partner Dr. Cristian Hofmann who motivated me a lot to finish this thesis and who supports my academic career in going to Vienna with me.

I am also very thankful for the hard work my student assistants and diploma and bachelor graduates did. They contributed a lot to this thesis and did studies who already go beyond the content of the here reported studies. Some of them are doctoral students, too, in the meantime. I especially thank the graduates Mara Gerich, Stefanie Jauch, Janna Hupp, Michael Kutscher, Caroline Müller, Alexandra Margic, Katharina Ernst, Judith Kramwinkel, Rebekka Kurzweil and Sven Beck as well as the student assistants Vlada Solovieva and Jessica Lang.

Last but not least many thanks go to the special interest group in teachers' competences' of the DFG SPP "competence models" for the interesting and helpful meetings full of ideas and support and to all the teachers and teacher students who took part in the studies.

Table of contents

VOLUME II:

Appendix

Summary

Diagnosing the abilities of students is one of the most central tasks that teachers need to perform in order to create an effective class that will meet the needs of their students. So far, accuracy in teachers' judgments has been measured by correlating their judgments with the results of standardized tests. However, an ongoing request exists to shift the focus from diagnosing students' achievements to diagnosing learning behavior in order to allow for didactic action afterwards. Furthermore, there is a call for further education programs in that field. The purpose of this dissertation was to develop and test a model of teachers' diagnostic competence that accounts for students' learning behavior. The model should close the gap between previous empirical research on diagnostic competence and recent theoretical demands. A further purpose of the study was to train teachers and teacher students in this new diagnostic competence by developing and evaluating a training program and a standardized diary based on the model. As a third purpose, the correlation between teachers' diagnostic competence and counseling competence should be tested. These competences obviously accompany each other, but the correlation has never been proved yet.

First, theoretical components of teachers' diagnostic competence were identified by summarizing multiple demands mentioned in the literature; thus building a process consisting of three postulated dimensions. Second, potential predictors of teachers' diagnostic competence that influence competence development were identified. Third and most importantly, the robustness of the claimed three-dimensional model was tested using confirmatory factor analysis and this model was compared with a g-factor model and a two-dimensional model. To test the model and measure diagnostic competence concerning learning behavior in the sense of the model, new instruments needed to be developed. Additionally, the validity of the newly developed scenario-test to measure diagnostic competence was considered. The influence of postulated predictors of competence

development was also tested for (a) a group of teachers, (b) teacher students in the first phase of German teacher education, and (c) teacher students in the second phase of teacher education. Finally, differences in the levels of diagnostic competence of teachers, teacher students in their second phase, and teacher students in their first phase of teacher education were tested.

Results indicate that the hypothesized three-dimensional process model indeed provides a very good and substantially better fit than the other models, and – for validation purposes - it is possible to predict an appropriate diagnosis by the model content. Knowledge and professional self-concept turn out to be substantial predictors of diagnostic competence, but reflected experience unexpectedly does not. Teachers and students in the first phase of education differ significantly in levels of competence, and students in the second phase differ from students in the first phase. However, there is no significant difference between teachers and students in the second phase. The first paper included in this dissertation is about the model, its predictors, the scenario-test and the differences in competence levels.

In the second paper, a newly developed training program and standardized diary based on the tested model, are dealt with. In that study, pre- and posttest measures were combined with time-series data to evaluate the training program. Results show that the training program does enhance teachers' diagnostic competence, especially when it comes to actions before and while diagnosing. The diary proves to be an accurate instrument to measure transfer, but it has no additional intervention effect to the training program. As the demand for diagnosing learning behavior and fostering students individually increases, the concept proves to be helpful both in teacher education and further education.

The third paper deals with the correlation between teachers' diagnostic competence and counseling competence, because diagnosing is especially relevant to counseling as it permits a teacher to provide accurate feedback and giving feedback in turn is part of the

postaction dimension of the tested model. Diagnostic- and counseling competence are measured within the same sample and a statistically significant correlation was found for total scores as well as for selected predictors of the competences. In latent regression analysis, counseling competence could be predicted significantly by diagnostic competence, but when the regression was computed separately for teachers, and two groups of teacher students, a prediction was only possible for teachers. Nevertheless, the data finally shows the correlation between diagnostic and counseling competence empirically.

Part 1 of this dissertation consists of a synopsis. It gives a theoretical introduction to the topic leading to the research aims, followed by an overview of the three papers and a summarizing discussion.

In part 2, the three original papers can be read, beginning with the one, in which the model is tested, followed by the paper about the training program and standardized diary and ending with the paper about the correlation of diagnostic and counseling competence.

As a result of this dissertation, a three-dimensional model of diagnostic competence that accounts for students' learning behavior has been established. It builds a profound basis for training programs and should be considered for correspondent modules in teacher education. With the developed scenario-test, an adequate instrument to measure the competence close to real behavior, but anyhow efficient, exists. The training program is effective in most variables and the standardized diary seems to be a promising instrument to measure the application of diagnostic strategies every day at school. Furthermore, the correlation between diagnostic and counseling competence could be shown empirically. This relation should be considered in teacher education and further training programs which could aim at fostering both competences with regard to their correlation.

Part 1: Synopsis

Introduction

Teachers are faced with multitasking highly complex work in their daily job routines (Brante, 2009). There is: having professional knowledge, giving learner-centered instruction, managing classrooms, interacting with students, and being motivated role models. In addition to that, diagnosing students' achievements and learning behaviors is one of their most central tasks (Kukla-Acevedo, 2009; Opdenakker & Van Damme, 2006).

Weinert (1999) defines competences as "a roughly specialized system of abilities, proficiencies, or individual dispositions to learn something successfully, to do something successfully, or to reach a specific goal" (p. 44). Applying his concept of key competences to the school setting, diagnostic competence is one of those key competences that teachers require (Weinert , 2001). It is of such an importance because "teachers are challenged to meet diverse learning needs and to adapt their teaching to heterogeneous academic ability as well as to multiple interests and motivations" (Vogt & Rogalla, 2009, p. 1051) in order to create an effective class that meets the needs of their students.

With regard to their research topic of language teaching, Edelenbos and Kubanek-German (2004) defined teachers' diagnostic competence as "the ability to interpret students' foreign language growth, to skillfully deal with assessment material and to provide students with appropriate help" (p. 260). This definition can be generalized for diagnosing learning behavior cross-curricularly as the ability to interpret students' academic growth and their growth in using learning strategies.

In prior empirical research diagnosing students' academic achievements was primarily dealt with. From the 1970s until today, teachers' diagnostic competence has been operationalized as their ability to accurately judge their students' achievements or task difficulties. So far, accuracy in diagnosing has been measured by correlating teachers' judgments with the results of standardized tests (e.g., Bates & Nettelbeck, 2001; Coladarci,

1986; Demaray & Elliot, 1998; Feinberg & Shapiro, 2003; Hecht & Greenfield, 2002; Helmke & Schrader, 1987; Lee, Chiu, van Hasselt, & Tong, 2009; McElvany et. al., in press; Meisinger, Bradley, Schwanenflugel, & Kuhn, 2010; Spinath, 2005; Wang, 1973).

However, there is an ongoing theoretical request to shift the focus from diagnosing students' achievements to diagnosing learning behavior in order to allow for didactic action afterwards (Abs, 2007). The aim is that diagnoses should not end in themselves, but rather enable teachers to foster their students individually and to adapt their classes to their students needs (e.g., Fend, 2006; Horstkemper, 2004; Kretschmann, 2009; Vogt & Rogalla, 2009; Winter, 2006). Nevertheless, accuracy in judgments is still an important component of teachers' diagnostic actions, but it should not be the sole representation of diagnostic competence anymore.

To close the gap between theoretical demands and empirical practice, the first research aim of this dissertation was to develop and test a model of teachers' diagnostic competence. For that, theoretical components of teachers' diagnostic competence were identified by summarizing multiple demands mentioned in the literature. Some theoretical assumptions, the chosen components and predictors of the model are now described.

To begin with, the model deals with teachers' diagnostic competence concerning students' learning behavior cross-curricularly, because diagnosing is necessary in every school subject and especially learning behavior and the application of learning strategies can be observed while learning different content. Context- and domain-specificity as an important attribute of competences is considered (Koeppen, Hartig, Klieme, & Leutner, 2008). The domain is diagnosing in the educational setting. Concerning context specificity, teachers' diagnoses of pupils' learning behavior both at school and at home are focused in particular.

Learning behavior addresses questions such as how pupils deal with tasks; how they do their homework; how they learn at home; which learning strategies they are able to apply,

and how much of a self-regulated learner they are. If problems occur, these issues of self-regulated learning can be fostered by the teacher (e.g., Perry, Hutchinson, & Thauberger, 2008; Perry, VandeKamp, Mercer, & Nordby, 2002).

Koeppen et al. (2008) also ask for theoretically sound and empirically tested competence models on which valid measures of competence need to be based: "These models have to (a) represent the internal structure of competencies in terms of specific basic skills and abilities, (b) describe different levels of competencies with reference to domain-specific performance, and (c) take into account changes occurring in learning and developmental processes" (Koeppen et al., 2008, p. 62). In addition to the theoretical framework of the model, the empirical testing should lead to a "methodological-substantive synergy" (Marsh & Hau, 2007). In this dissertation, the characteristics of such competence models are implemented as follows: The internal structure is specified as three-dimensional. Following the terminology of models of self-regulation (Schmitz & Wiese, 2006; Zimmerman, 2000), the dimensions are called preaction, action, and postaction phase. If there are different levels of diagnostic competence in a teachers' career will be tested by a comparison of three cohorts with different job experience. To take into account changes in teachers' developmental processes, potential predictors of diagnostic competence will be tested for their predictive power.

Last but not least, the model is conceptualized as a process (Jäger, 2007), again comparable to models of self-regulated learning (Schmitz & Wiese, 2006; Zimmerman, 2000). Figure 1 illustrates the process model. The three dimensions will now be further described.

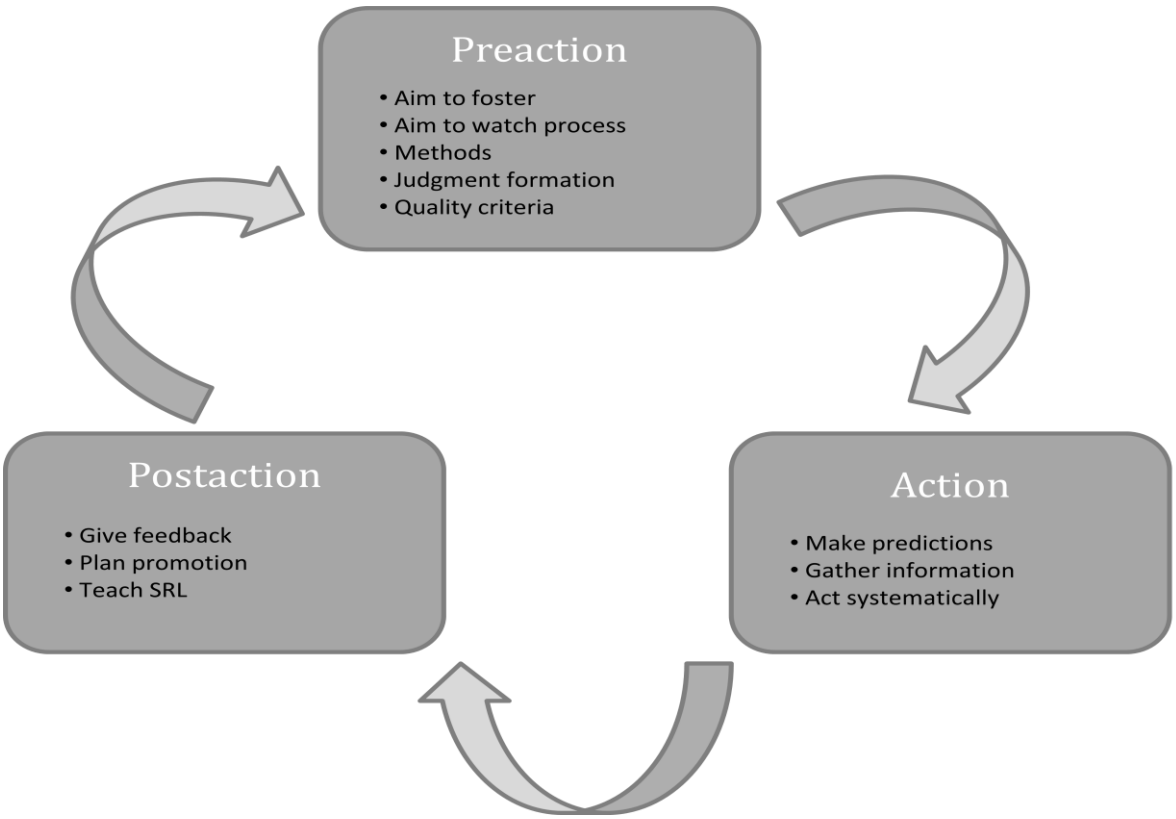


Figure 1. Process model of teachers' diagnostic competence concerning pupils' learning behavior.

The first dimension is called *preaction phase*. In that phase, every diagnostic action before actually diagnosing matters. It is especially important that the teacher sets the aim of the diagnosis, which means that the teacher should intentionally aim to watch the individual student's learning process and to foster the student based on the diagnosis (Abs, 2007; Horstkemper, 2004; Kretschmann, 2009). The teacher should develop an individual frame of reference to reduce the big-fish-little-pond effect (Lüdtke, Köller, Marsh, & Trautwein, 2005). Additionally, the teacher's basic diagnostic skills are activated (Strasser & Gruber, 2003), composed of knowledge about methods for gathering information (Arnold, 1999; Helmke, Hosenfeld, & Schrader, 2004), knowledge about psychological quality criteria of tests, and knowledge about judgment formation (Ophuysen, 2006). At best, the teacher should not only be familiar with, but also know how to deal with these methods, and know the situations in which each method is the most effective. Following Maclellan (2004), "it is not sufficient that

teachers be procedurally skilled in implementing assessment (necessary as such skills are) but they must also be able to reason about their assessment practice" (p. 524). Her study showed that teachers "do not connect issues of replicability and generalisability (and therefore of the issues of reliability and validity) with assessment methods" (p. 530). Knowledge about judgment formation is a construct taken from social psychology where systematic biases due to judgment heuristics influence a person's decision making. Within classrooms, such biases can skew teachers' judgments, too. Fiedler, Walther, Freytag, and Plessner (2002) investigated systematic biases in teachers' judgments of student achievement in a simulated classroom. If a teacher wants to come to an adequate diagnosis, it is necessary to know theses biases and be aware of them in order to be able to avoid them.

The second dimension of the model is called *action phase*. In that phase, the actual diagnosis is made. Acting systematically is most important here. The systematic approach is based upon proceeding scientifically in doing quantitative research (Wilson, 1952), beginning with making a prediction about a student's development and possible underlying learning difficulties. Making predictions while diagnosing is adopted from medicine where clinicians can use clinical prediction rules "to predict the most likely diagnosis, prognosis, or response to treatment in a patient based on individual characteristics" (McGinn, Jervis, Wisnivesky, Keitz, & Wyer, 2008, p. 1261). To make a prediction, the teacher has to gather information from different sources and choose the relevant ones. In the end, the teacher can interpret the data and come to a concluding diagnosis. By comparing real developments with the ones he predicted it is possible to initiate changes in his modus operandi for his next diagnosis.

The third dimension of the model is called *postaction phase*. It begins right after a diagnosis has been made and consists of the demanded pedagogical action afterwards (Abs, 2007; Kretschmann, 2009). Giving feedback to students is important here. Feedback "is among the most critical influences on student learning" (Hattie and Timperley, 2007, p. 102).

In their conceptual analysis of feedback, the authors give advice about the conditions under which feedback is the most effective. Likewise, Butler and Winne (1995) stated that "feedback is inherent in and a prime determiner of processes that constitute self-regulated learning" (p. 245). Feedback to parents about the diagnosis is also important and can be seen as a main element of counseling parents. Diagnosing and counseling are strongly connected. In a model of counseling competence by Bruder, Klug, Hertel, Kelava, and Schmitz (under revision), diagnosing turns out to be one of four dimensions aside from counseling skills, cooperation/perspective taking, and coping. The second important content of the postaction phase is writing down plans for the individual student's promotion. Teachers do have to write such plans in Germany if a student is at-risk. In these plans, (a) the characteristics of the student that should be fostered, for example, motivation, concentration, creativity, ability to think abstractly, and so forth; (b) the student's actual skill level; (c) the goals to be reached; and (d) the measures that will be used to reach these goals should be written down. The third postaction content concerns adapting the class as a reaction to the diagnosis by means of teaching appropriate learning strategies and self-regulated learning (SRL). Several studies have emphasized the relevance of self-regulated learning for students' academic achievement. Pintrich and van de Groot (1990), for example, found that "self-regulation was the best predictor of academic performance […,] which suggests that the use of self-regulating strategies, such as comprehension monitoring, goal setting, planning, effort management and persistence, is essential for academic performance on different types of actual classroom tasks" (p. 38). A study also showed that "it is possible to support self-regulation competencies and mathematical achievement by self-regulation intervention within regular mathematics lessons" (p. 17) held by teachers (Perels, Dignath, & Schmitz, 2009).

As mentioned before, diagnostic competence is conceptualized as a process. Thus, the model is of a cyclical nature, in which the three dimensions can influence each other, and in

particular, there is a connection between the postaction phase in one diagnosis situation and the preaction phase in a consecutive diagnosis situation.

After having described the model theoretically, the potential predictors of diagnostic competence will now be further illuminated. The predictors can give insight in which variables are crucial to further develop diagnostic competence. Potential predictors for diagnostic competence are chosen with reference to medical diagnosing, teacher professionalization and expertise research. In the field of medical diagnosing, Epstein and Hundert (2002) state that competence builds on a foundation of basic clinical skills, scientific knowledge, and moral development. A competent clinician needs to acquire and use knowledge and to integrate data in clinical reasoning to solve real-life problems. Furthermore, he needs the willingness, patience, and emotional awareness to use these skills humanely. Bakkenes, Vermunt, and Wubbles (2010) recently defined teacher learning as an active process in which teachers engage in activities that lead to a change in knowledge and beliefs and/or teaching practices. Strasser and Gruber (2003) state that expertise in the field of counseling is build on knowledge, reflected experience, and personal resources. The so-called COACTIV model developed by Kunter et al. (2007) illustrates teachers' knowledge and beliefs as two key components of teachers' expertise. Shulman and Shulman (2004) integrated personal variables such as vision and motivation and cognitive variables such as understanding, practice, and reflection into their model of teaching. In particular, reflection on experienced action, which is addressed as an additional deliberate practice, seems to be a key factor for competence development (e.g., Berliner, 2001). The central idea is that "through reflection the teacher better understands and extends his/her professional activity, and that reflecting on teaching problems will lead to new insights for practice" (Marcos, Miguell, & Tillema, 2009, p. 191). It allows practitioners to examine their own clinical reasoning strategies (Epstein & Hundert, 2002). In Clarke and Hollingsworth's (2002) model of

teachers' professional growth, the symbiosis of reflection and enactment is the central mechanism for gaining professional growth, and Sowa (2009) was able to show that reflection was a crucial factor that contributed to changes in teaching. Which are the main variables all these models have in common? It is mainly three things: knowledge, personal variables and reflection on own experienced action. Thus, applied to teachers' diagnostic competence, the three chosen variables to serve as predictors are: (a) knowledge about diagnostics, (b) reflected experience in the field of diagnosing, and (c) professional self-concept in diagnosing, which consists of certain attitude variables (motivation to diagnose, attitude toward diagnosing, self-efficacy in diagnosing, and conscientiousness) and serve as personal resources the teacher brings with him.

The assumptions about the model as an advancement of prior research on diagnostic competence, its character, dimensions and predictors served as the theoretical basis for each of the studies included in this dissertation. After its fit has been tested and compared to competing models, it could build a profound basis for training programs in diagnostic competence. Following Bakkenes et al. (2010), there is a growing awareness of the necessity of assisting teachers in their professional development in general. In particular, there is a call for further education programs to foster new facets of teachers' diagnostic competence (Klieme et al., 2003). Nevertheless, there are few such programs yet. Therefore, the second research aim of this dissertation was to develop and evaluate a training program to foster teachers' diagnostic competence concerning learning behavior based on the model. In the training program, the three phases of the model are incorporated. Additionally, teachers' knowledge, reflected experience and professional self-concept in diagnosing should be fostered simultaneously within the training program. To facilitate reflecting on the training content, a standardized diary was additionally implemented for a group of teachers. This diary contains questions on every phase of the diagnostic process. According to Webber,

Scheuermann, McCall and Coleman (1993) the continuous registration of one's learning behavior can lead to personal modifications of behavior in a desired direction. Schmitz and Perels (2011) could prove, that students who work on a learning diary, come to better results in mathematical problem solving, self-regulation and self-efficacy than a control group. Thus, the assumption was that self-monitoring diagnostic behavior could also lead to a modification of teachers' diagnostic action and self-concept in a desired direction, so that they can further enhance their diagnostic competence and better transfer the training content to the classroom. With the help of the process data that can be gained by the diaries, the application of the learned strategies in class can be measured and analyzed in a sophisticated way.

Not only diagnosing learning behavior is an important competence in the teaching profession. Looking at the postaction phase, giving feedback, including counseling parents, is a crucial variable. Teachers' counseling competence is just as well one of teachers' key tasks (KMK, 2004). Knowledge about counseling is also a component in models of teacher professionalization (e.g. Baumert & Kunter, 2006). If parents request counseling, they do not only consider teachers with a special education in counseling. In fact, every teacher has to do counseling talks (Landesinstitut für Schule und Weiterbildung, 1998). Furthermore, there is a growing demand for counseling at school, in particular for counseling in learning strategies (Schnebel, 2007). Parents do desire teachers' counseling, because they feel insecure in how to support their child in learning (Wild, 2003). However, 94% of teachers do not feel well prepared for counseling parents by their education (Hertel, 2009). Furthermore, they often feel overstrained while counseling (Hitzinger, 1987). Particularly those teachers, who think that there are not enough further education programs in counseling, cooperate less with parents (Wild, 2003). Studies further show the importance of an intense cooperation between parents and teachers (e.g. Epstein & van Voorhis, 2001). Cox (2005) could show in a meta-analysis that interventions to ameliorate the cooperation between parents and teachers are effective and

do improve both students' academic achievement and behavior at school. Current studies aim at fostering teachers' communication and negotiation skills with the help of training programs (Aich, 2006; Hertel, 2009). In a model of teachers' counseling competence by Bruder (2011), which focuses the domain of counseling concerning learning strategies, diagnosing builds one of four dimensions. Given that counseling is integrated in the postaction phase of the described model of diagnostic competence, too, there seems to be a shared part of variance. McLeod (2003) sees diagnosing as a basis for counseling. It is evident that a profound diagnosis is useful to counsel adequately, but that does not mean that a good diagnostician must in turn be a good counselor. Nevertheless, there is an evident relation that has never been tested empirically. So far, teachers' competences in general, as a rather new research topic, were just investigated separately. However, their interaction can deliver new insights for competence research and teacher education.

Thus, the third research aim of this dissertation was to test the correlation of diagnostic competence and counseling competence in order to verify the evident relation empirically. The testing should be controlled for the influence of the measurement method. Furthermore, the predictors in the models of both competences are similar except for the respective domain. Thus, they are tested for a substantial correlation, too. If the correlation can be shown, it would be indicated to develop combined training programs that cover both competences and to teach both symbiotically in teacher education.

Overview of the manuscripts

In this section, an overview of the three manuscripts contained in this dissertation is given by summarizing the method and main results of each manuscript before the results of all the studies will be discussed together. One manuscript is under review, one under revision and one is in press in a pertinent scientific journal listed in the Social Science Citation Index (SSCI).

Summary of manuscript 1.

Manuscript 1 is about the development and testing of the model of teachers' diagnostic competence concerning learning behavior described in the introduction. Furthermore, the instrument to measure diagnostic competence is validated by the accuracy of the teachers' diagnosis within the case-scenario, the three predictors deduced from research on teachers' professionalization and expertise research are tested for their predictive power and teachers' at different stages of their career are tested for differences in their level of competence. Thus, the first research aim is pursued in manuscript 1.

A sample of N = 293 teachers and teacher students (93 grammar school teachers, 107 teacher students in their second phase of teacher education, and 93 teacher students in their first phase of teacher education) participated in the investigation. Participants completed three tests. Diagnostic competence concerning learning behavior was measured by a scenario test with open questions based on the model. Knowledge of diagnostics as a predictor was measured by a multiple-choice knowledge test and the two other predictors, professional self-concept and reflected experience, were measured by means of a questionnaire via self-assessment. Participation was voluntary, and each participant received an incentive.

Confirmatory factor analysis showed that the postulated three-dimensional model with correlated factors fits the data very well with a non-significant chi-square test and a SRMR and RMSEA much lower and a CFI higher than the cutoff-criteria for fitting models. Compared to a one- and a two-dimensional model, the three-dimensional model shows a better fit with an AIC and BIC of the three-dimensional model that are smaller than the ones of the other models. Furthermore, the chi-square difference test indicated that the models differ significantly.

Concerning validation, multiple regression analysis with the dimension scores from the case scenario as predictors and the correctness of the diagnosis measured by a validation

question in the case scenario as the criterion showed that it is possible to significantly predict the correctness of the diagnosis from the teacher's preactional, actional, and postactional dimension scores.

To test the predictive power of the assumed predictors, multiple regression analyses were calculated separately for the three participating subgroups. Results revealed that in the group of teachers professional self-concept is the best predictor of diagnostic competence. This means that the more motivated, interested, self-efficient and conscientious teachers are in diagnosing, the better the higher is their diagnostic competence and vice versa. Knowledge can also significantly predict diagnostic competence, but only for the postaction dimension. In the groups of teacher students in their second phase of teacher education and teacher students in their first phase of teacher education only knowledge could significantly predict dimensions of diagnostic competence. Unexpectedly, reflected experience appeared to be no substantial predictor, except for a negative prediction of the postaction dimension in the group of teachers that was just as well not expected.

Using ANOVA with group as the independent variable and diagnostic competence as dependent variable, the level of competence in the three subgroups was compared. Teachers and students in the first phase of education differed significantly in levels of competence, and students in the second phase differed from students in the first phase. Teachers and students in the second phase showed higher values than students in the first phase of German teacher education. However, there was no significant difference between teachers and students in the second phase.

As a conclusion, the construct of teachers' diagnostic competence has been broadened with the empirically tested and validated process model. Teachers' knowledge and professional self-concept have been shown as important for competence development. Furthermore, a suitable instrument exists for measuring that competence. On that basis, an

amelioration of teacher education and further education can be implemented. In particular, training programs for teachers who are already on the job are needed. In addition to the implementation of training programs, reflected experience, which was insufficiently measured by self-assessment in this study and probably therefore did not prove to be a substantial predictor, can be stimulated and manipulated in a more adequate way by the help of diaries in order to aid teachers' diagnostic competence.

Summary of manuscript 2.

Manuscript 2 is about the evaluation of the training program and the standardized diary and covers the second research aim. It conforms to the demand for further education programs that is discussed in manuscript 1. The empirically tested process model of teachers' diagnostic competence concerning learning behavior builds the theoretical basis of the training program and the standardized diary. In the training program, the three dimensions of the model (preaction phase, action phase, postaction phase) are incorporated. Not only the variables of diagnostic competence contented in the model, but also the tested predictors were considered in the intervention. Furthermore, reflected experience should be manipulated by the application of standardized diaries which supplies the transfer of the learned content by means of self-monitoring. The diary contains questions on every variable of the diagnostic process. In contrast to a control group, an increase in teachers' diagnostic competence concerning learning behavior and teachers' knowledge, reflected experience and professional self-concept was expected if they participated in the training program. For the teachers who work on the diary, an additional intervention effect to the one of the training program was expected because of the supplementary self-monitoring. In the process data collected by diaries, positive linear trends were expected for each trained variable as well as for reflected experience and professional self-concept over the training period. Furthermore, for the diary

data, a lasting augmentation of scores from a baseline for each training variable just after the session in which the specific variable was trained, was expected.

47 grammar school teachers participated. Participation was voluntary and teachers got an incentive. The longitudinal quasi-experimental design combined pre- and posttest measures with time-series data. For pre- and posttest, a multi-method approach was chosen consisting of the three instruments which have already been used for the testing of the model: (a) a scenario test with open questions to measure diagnostic competence based on the model, (b) a multiple-choice knowledge test to measure knowledge in diagnostics and (c) a questionnaire to measure professional self-concept and reflected experience in diagnostics via self-assessment. Concerning the design, there were two experimental groups who participated in the training and one control group, who participated in pre- and posttest and was offered the possibility to get a shortened training program afterwards. Experimental group 2 additionally worked on the standardized diary, starting one week before the first training session and finishing one week after the last session. The process data were gained by the standardized diagnosis diaries teachers in the experimental group 2 got, which were newly constructed with reference to the diagnostic process. Each diary item is formulated as a state in contrast to questionnaires which usually survey traits.

The training program covers the three phases of the diagnostic process. In each session a great extent of activity and reflection was realized. Participants worked on a specific own case of a student. Various methods and social forms were used to learn and practice the content of the diagnostic progress. Participants had to do homework additionally.

Data of the pre- and post- test were analyzed using a multivariate one-way ANOVA with group as independent variable and the pretest- posttest differences of the scenario test, knowledge test and questionnaire measures as dependant variables. Results show that the training program does enhance teachers' diagnostic competence especially when it comes to

actions before and while diagnosing. There are significant differences in the pre-post-difference of the groups for nearly each dependant variable with small to medium effect sizes. Not only diagnostic competence, but also knowledge, professional self-concept and reflected experience increased. Contrasts revealed that, as expected, both experimental groups had a significantly higher increase than the control group, but the increase of EG 1 and EG 2 does not differ significantly.

The return rate of the diaries was 59% (176 analyzable diaries). Trend analyses showed significant linear trends for most of the diary variables. In the course of the training program and the work on the diary, teachers did more and more apply the learned strategies in class still one week after the last training session took place. With the help of interrupted time-series analyses, we were able to analyze the effect of each trained variable just after the training session in which the particular content was taught. Results showed that there was an intervention effect in diary data that stayed stable or further augmented for most of the preaction and action variables of diagnostic competence.

Summing up, the training program turned out to be effective in promoting teachers' diagnostic competence concerning learning behavior. The diary proved to be an accurate instrument to measure transfer, but it had no additional intervention effect to the training program. As the demand for diagnosing learning behavior and fostering students individually increased, the training concept seems to be helpful both in teacher education and further education.

Summary of manuscript 3.

Manuscript 3 is about testing the correlation between teachers' diagnostic competence and counseling competence empirically and thus deals with the third research aim of this dissertation. The testing was controlled for the influence of the measurement method and the

predictors in the models of both competences are tested for a substantial correlation additionally.

There obviously seems to be a relation between both competences that has never been verified empirically yet. The basis built the model of teachers' counseling competence from Bruder (2011) and the described model of teachers' diagnostic competence concerning learning behavior. In both models, the respectively other competence is integrated as parts of one dimension.

Within the same sample as in manuscript 1 consisting of N = 293 teachers and teacher students (93 grammar school teachers, 107 teacher students in their second phase of teacher education, and 93 teacher students in their first phase of teacher education), diagnostic and counseling competence were measured using both case scenarios, one scenario followed by open questions with reference to the according model for each competence. Counseling competence was additionally measured using a situational judgment test in order to control the influence of the measurement method. Furthermore, knowledge in both competences was measured using tests in a multiple-choice format and reflected experience and professional self-concept in both competences were measured using questionnaire items.

We expected to find correlations between diagnostic competence and counseling competence measured with the equivalent instrument and between the predictors of both competences of a medium size. If both competences were measured with non-equivalent instruments, we expected a small correlation. In a continuative analysis, we tested whether there are differences in the prediction if it is calculated separated for the three groups (teachers, teacher students in the second and teacher students in the first phase).

Results show that the expected correlation occurs significantly between diagnostic competence and counseling competence and their predictors measured with an equivalent instrument. Nevertheless, rather small correlations occur than the expected medium ones. If

the competences are measured with non-equivalent methods, the correlation decreases indeed, but still significantly occurs. Latent regression analysis to predict the dimensions of counseling competence by the dimensions of diagnostic competence shows similar results with very good fit indices for the model and a regression coefficient a little bigger than the correlations. When we computed a manifest regression separately for the 3 groups in the continuative analysis, we could only make a significant prediction for the group of teachers. Even though the correlative analyses do not allow for causal interpretations, we assume that a solid diagnosis precedes a good counseling session. The data shows empirically that the postulated relation between diagnostic and counseling competence exists. It would be helpful to teach this relation in teacher education and to develop training programs that cover both competences in their interdependency.

Summarizing discussion

Results of manuscript 1, which covered research aim 1, indicate that the newly developed three-dimensional model indeed provides a good fit. It fits substantially better than a one- or two-dimensional model. Although the one- and two-dimensional models are more economical, the three-dimensional structure fits the empirical data best. Thus, the new concept of diagnostic competence seems to be a multidimensional construct, even if the three dimensions are substantially correlated. The correlations emphasize the postulated process character of diagnosing consistent with Jäger (2007). Furthermore, it was possible to predict appropriate diagnoses from teachers' diagnostic competence test scores that were based on the model. Thus, teachers who follow the steps of the diagnostic process advised by the model in fact come to a more adequate diagnosis about students' learning behavior and vice versa. This can be seen as a validation of the variables comprised by the model. Two of three postulated predictor variables proved to be substantial predictors for diagnostic competence. In the group of teachers, their professional self-concept was the best predictor followed by knowledge

about diagnostics. The more motivated, interested, self-efficient, and conscientious the teacher is in diagnosing and the more knowledge he has about diagnostics, the better he is in preactional, actional, and postactional diagnostic competence and vice versa. In the other two groups of teacher students in the first and second phases of teacher education, knowledge about diagnostics was the only relevant predictor of their diagnostic competence. Professional self-concept obviously does not become relevant for competence development until the teacher is finally on the job, whereas knowledge is already important in the beginning of a teachers' career and still is in a later stadium. The finding that cognitive variables such as knowledge about diagnostics and personal variables such as professional self-concept are relevant for the development of competence are in line with findings from expertise research (e.g., Bakkenes, Vermunt, & Wubbles, 2010; Epstein & Hundert, 2002; Shulman & Shulman, 2004; Strasser & Gruber, 2003) and the COACTIV model (Kunter et al., 2007). Reflected experience unexpectedly did not appear to be a significant predictor in the student groups and even a negative one in the group of teachers. This does not support literature on reflected experience, where reflected experience has usually been assumed to be a crucial predictor of competence (Berliner, 2001; Bruder, Klug, Hertel, Kelava & Schmitz, submitted; Epstein & Hundert, 2002; Clarke & Hollingsworth, 2002; Marcus, Miguell, & Tillema, 2009; Strasser & Gruber, 2003; Sowa, 2009). A possible explanation for the small negative correlation could be ruminative behavior instead of the desired self-reflection. Teachers who state that they reflect a lot on their diagnostic actions probably do not come to an action in the diagnostic process and vice versa. Trapnell and Campbell (1999) and Grant, Franklin, and Langford (2002) for example give evidence for ruminative self-reflection. Besides, reflected experience was measured by self-reports in this study, which are prone to social desirability. Teachers estimated their reflected experience with $M = 4.30$ ($SD = 0.97$) on a six point Likert-scale rather high with low variability. A more objective way of measuring reflected experience

could lead to other results. Thus, a manipulation of reflected experience by implementing standardized diaries which should stimulate teachers to self-monitor their diagnostic action was planned for the training study in manuscript 2. Within the diaries, process data about reflection should be collected with one item that asks for the daily reflection on diagnostic action and the intervention effect of the additional reflection stimulated by the whole diary should be tested. The results of the comparison of teachers' competence levels at different stages in their career confirmed the hypotheses in part. Teachers with professional experience were expected to be more competent diagnosticians than students in the second phase of teacher education who in turn were expected to be more competent than students in the first phase. As postulated, there was a highly significant difference between teachers and teacher students in the first phase of teacher education and between teacher students in the second phase and teacher students in the first phase. Teachers and teacher students in the second phase showed higher values in diagnostic competence than students in the first phase of teacher education. Nevertheless, there was no significant difference between teachers and teacher students in the second phase of teacher education. Bruder (2011) found a similar pattern when comparing the three groups' counseling competence. A different level of motivation in working on the instruments does not seem to explain the absent difference in competence level between teachers and teacher students in the second phase of their education. Motivation was measured and groups were tested for differences, but there was none. Nevertheless, motivation was measured by self-report, which may be biased due to social desirability. An explanation for the comparatively good result of teacher students in their second phase of teacher education can be found in a change in German teacher education a few years ago when the second phase was modularized (KMK, 2000). Now, some modules exist that broach the issue of diagnosing, fostering, and counseling, and these seem to be very effective. Nevertheless, teachers on the job who did not get that kind of education lack

knowledge about the new content and perhaps compensate for that lack of knowledge with a more professional self-concept, which in the end leads to comparable values in diagnostic competence. Descriptive analyses revealed that there is also a need for change in the first phase of teacher education. Participants state that diagnosing of learning behavior has rarely been addressed in that phase (M = 1.90, SD = 1.31). Consequently, it is not astonishing that students in the first phase do not perform very well. There is a high need for fostering diagnostic competence across all groups. Out of 36 possible points in the scenario-test M = 14.28 (SD = 4.59) were averagely gained with a maximum of 26 points and a minimum of 0 points. Thus, there is much leeway left to arrive at an optimal result. Consequently, a training program as it is implemented in manuscript 2 to foster teachers diagnostic competence seems necessary. The instruments developed and used in this study proved to be appropriate by showing satisfactory quality criteria such as high inter-rater reliabilities for the scenario test, medium item difficulties for the knowledge test, and acceptable internal consistencies for the questionnaire. The scenario test, which was used to measure diagnostic competence concerning learning behavior, is both close to measuring real teacher behavior while still being economical. It sure does not measure real action, but it approaches the action the teacher would show in a real situation. There is just the unsolved problem of whether reflected experience can be measured more accurately in any way other than self-assessment. Despite this, the instruments proved to be good enough to be used in further studies, as it is done in manuscript 2 where they are applied to measure changes after an intervention in a pre post control group design. The greatest limitations of manuscript 1 are definitely the voluntary participation and the cross-sectional design. Because of the voluntary participation the sample is not representative. The assumption stands to reason that the participants are the rather motivated ones out of the population of teachers and maybe even the ones with rather high diagnostic competence. Nevertheless, the values even of this sample were rather low to

medium on the scenario test. It is to be assumed that values in the population are even lower than in the investigated sample.

To summarize the findings of manuscript 1, diagnostic competence concerning learning behavior seems to be a multidimensional process that can be measured and fostered based on the tested three-dimensional model. Based on the tested model a training program should be implemented as a first step before curricula and modules in teacher education on this topic should be developed or modified in further studies. Furthermore, the operationalization of reflected experience is tried to be optimized. The idea is to manipulate reflected experience in the training study by the application of a standardized diary in which teachers are asked to reflect on their diagnosing experiences. Simultaneously, the diary should promote the transfer of training contents into the classroom, which thereby could be measured.

Such a training study and a standardized diary were dealt with in manuscript 2. The results of the pretest- post-test comparison show that, as expected, teachers do benefit from the training program in contrast to a control group. Diagnostic competence concerning learning behavior increased in both training groups in the overall score and the preactional and actional dimension. Not only competence increased, but also knowledge in diagnostics showed an increase of the same pattern. In addition to that, teachers' professional self-concept and reflected experience in diagnostics increased if they took part in the training program, even if those two variables were not explicitly covered in the training program. As knowledge and professional self-concept are relevant predictors for the development of diagnostic competence in manuscript 1, it is a great advantage if they augment along the way. Reflection of experiences also is seen as an important condition for the development of competences in expertise research (e.g. Strasser & Gruber, 2003) even if it was no significant predictor of diagnostic competence in manuscript 1. Thus, participation in the training program not only

increases teachers' diagnostic competence directly, but also allows for getting a basis to further develop their competence. Effect sizes were all small to medium, but in a range that is meaningful, considering the short duration of the training program with only three 180-min sessions. Unexpectedly there was no significant effect due to the training program for the postaction measures of diagnostic competence in pretest- post-test comparison, but there was at least a tendency in the expected direction. In a training program of a longer duration in which the complex postaction content can be trained in more detail, a significant effect should occur. Planning promotion of students, giving feedback to students and parents and teaching self-regulated learning is far too much content for the one short training session in the present study. Concerning the diary, the return rate of 59% can be considered as a great success. Teachers accepted the short standardized form of the diary and worked on it constantly over the four weeks. The standardization of the items and the length of just one page seem to be relevant to motivate teachers to work on it. If it was longer and would cost more effort, the return rate probably would be lower. As a further source of motivation, teachers who worked on at least 80% of their diaries had the possibility to get extra credit points. The use of incentives seems to be a further important way to ensure participants commitment. Nevertheless, the short and standardized version is good for measurement, analyses and motivation, but it is at the expense of teachers' reflection and the desired additional intervention effect. The expected additional intervention effect did not occur. A possible explanation may be that there already was much reflection integrated in and stimulated by the training program, so that teachers who did not work on the diary also benefitted of self-monitoring. Maybe it was just expected too much of that short and standardized diary version in having an additional intervention effect to the training. However, the standardized diary proved to be a helpful instrument to measure the transfer of the trained variables to teachers' everyday work at school. With the diary, teachers did reflect on which training content they

really implemented in their class every day. Furthermore, the process data collected by the diaries give much more insight in which diagnostic action the teacher applies over time and how his augmentation in competence takes course. The positive linear trends for the trained variables as well as for teachers' professional self-concept and reflected experience that were expected over the four weeks could be found for eleven of the sixteen measured variables. There was even a significant linear trend for the postaction variable *plan promotion*, which did not occur in the pretest-post-test comparison, even when analyzed on item level. Thus, the diary data provides us with additional information. So do the interrupted time-series analyses. They allow illustrating and proving an augmentation of the scores from a baseline just after the session in which the specific variable was trained for a great extent of the trained variables. Furthermore, it can be seen if the augmentation stays, declines or even further rises. For the five significant variables, it does further rise, so that a long-term and even delayed effect of the training program and the work on the diary is proved. Similar to pre-post analyses, postaction variables showed no intervention effect in interrupted time series analyses. In addition to the needed extension of the training program especially when it comes to postaction content, the operationalization of the postaction variables in the diary needs to be improved. Teachers do not and cannot implement these variables (plan promotion, give feedback, teach SRL) every day at school. Thus, the formulation of the items, which for example ask for if the teacher did teach SRL that day, cannot be answered positively every day even if the intervention was effective. The postaction diary items need to be reformulated maybe asking for if the teacher did give thought to those variables that day.

The study design in manuscript 2 was a quasi-experimental one with the advantage of being longitudinal and combining pre- and post-measures with process data and having a control-group. However, the disadvantage is that in the field no randomized assignment of the teachers to the conditions was possible, because the teachers were trained in their schools and

had limited time for the appointments. But it was randomly assigned which group would be in the experimental and which in the control condition. Furthermore, in each of the three training groups half of the participants did work on the diary. All things considered, the quasi-experimental design fits the field research that was intended. As a sample, grammar school and comprehensive school teachers with a grammar school track were chosen similarly to the sample in manuscript 1. A generalization of the training program for other school forms should be intended.

Manuscript 3 was about testing the correlation between teachers' diagnostic competence and counseling competence empirically and thus dealt with the third research aim of this dissertation. Results revealed that the expected correlation between teachers' diagnostic competence and counseling competence could be shown indeed. Nevertheless, the correlation was rather small than medium, as hypothesized before, when measured by a similar method. A prediction of counseling competence by diagnostic competence was possible using latent regression analysis. When measured with non-similar instruments the correlation still significantly exists, but it becomes smaller. These results fit the hypotheses in terms of the multitrait-multimethod approach, because the part of shared variance that is due to the measurement method is controlled. That the correlation reveals to be small instead of medium seems plausible, because the competences differ in their domain. They just share a small part. In terms of the models it is just a part of one dimension in each model that covers the respectively other competence. It was too much to expect a medium correlation in the beginning. Concerning the expected correlations between the predictors of the two competences, a similar pattern exists. Professional self-concept, reflected experience and knowledge each with competences do show small but significant correlations. It can be concluded that teachers who have a lot of knowledge, reflected experience and a positive professional self-concept in one of the competences, they also have it in the other

competence. However, the corresponding predictors of the competences were measured with similar methods. Thus, a part of the shared variance can be due to the measurement method. Yet, you can suppose that the pattern would be the same as it is for the competences when non-similar methods are used. Hence, the correlation would be smaller but would still exist in a significant way. The continuative analysis in which the regression was calculated as a multi-group comparison had to be on a manifest with the overall values of both competences. An analysis on latent level should have been the method of choice, but it was not possible to calculate because of the number of subjects, which was too small. The multi-group comparison on manifest level showed that counseling competence could only be predicted significantly by diagnostic competence in the group of teachers, but not in the two groups of teacher students. The regression coefficient is bigger just for teachers than in the whole group of all participants and with 27% shared variance a significant part of teachers' counseling competence variance can be explained by their diagnostic competence. The question stays open why the correlation just occurs in the groups of teachers but not in the groups of teacher students. Looking at the achieved levels and standard deviations of the three groups in manuscript 1, teachers and teacher students in the second phase do differ from teacher students in the first phase, but teachers and teacher students in the second phase do reach similar values in their diagnostic competence. The same pattern occurs for counseling competence. The standard deviations do not differ in a substantial amount for both competences in the three groups. Thus, differences in levels or standard deviations of the three groups cannot suffice as an explanation for the correlation that just occurs in the group of teachers. Nevertheless, the correlation seems to occur only if the teacher is more experienced. Even if teacher students in the second phase do reach a similar level in both competences as teachers do, there is no systematic co-variation between the values of the two competences at their stage of professional development. Teacher students in the second phase

are rather good at one of the competences while the other is less far developed. However, teachers who are good diagnosticians predominantly seem to be good counselors, too, and vice versa. By looking at single cases, it becomes obvious that the difference between the reached values in diagnostic competence and counseling competence is often bigger in the group of teacher students in the second phase than in the group of teachers. Choosing a criterion of six points difference in the overall scores of the competences, only 12.2% of the teachers (n = 11) show this difference, whereas 35.9 % (n = 37) in the group of teacher students in the second phase and 34.4 % (n = 32) in the group of teacher students in the first phase have a difference in the competence values of at least six points. Furthermore, it is noticeable, that the group of teacher students in the second phase is quite heterogeneous containing students from the first until the fourth semester of the second phase. Particularly the students in the fourth semester (n = 8) are responsible for the reached competence level that does not differ from the one teachers reach.

As the data is gathered within the sample as in manuscript one, the same limitations occur here. Participation was voluntary and the design is a cross-sectional one. The assumption could arise that the participants are the rather motivated and competent ones, but reached values in both scenario-tests with which diagnostic and counseling competence were measured are rather moderate considering the possible values that could have been reached (M = 14.57, SD = 4.33, of 29 possible points for counseling; M = 14.28, SD = 4.59, of 36 possible points for diagnosing).

As a conclusion, the obvious and in the models assumed relation between diagnostic competence and counseling competence could be shown empirically in this manuscript. Nevertheless, the correlative analyses do not allow for a causal interpretation. It seems plausible that diagnostic competence is a necessary but not sufficient constraint for counseling competence, in particular because diagnosing should precede counseling in time. But this

assumption cannot be proved with the kind of analyses in this manuscript. However, diagnosing and counseling should be taught in a combined way which emphasizes their connection in teacher education and further education.

Limitations.

The greatest limitations are due to the cross-sectional design in manuscript 1 and 3 and the voluntary participation in all three manuscripts. Because of the voluntary participation, the sample is not representative. The assumption stands to reason that the participants are the rather motivated ones out of the population of teachers and maybe even the ones with rather high diagnostic and counseling competence. Nevertheless, the values in both competences even of this sample were rather low to medium on the scenario tests. It is to be assumed that values in the population are even lower than in the investigated sample.

Besides, reflected experience was measured by self-reports in manuscript 1 with a questionnaire and in manuscript 2 with a diary item. Self-reports are prone to social desirability and teachers estimated their reflected experience rather high with low variability in the questionnaire in manuscript 1. This could be the reason why reflected experience unexpectedly appeared not to be a significant predictor of diagnostic competence. Maybe other results could be gained when a more objective way of measuring reflected experience would be applied. However, reflected experience should be manipulated by the use of a diary in manuscript 2. The expected intervention effect that goes beyond the training program could not be shown. An explanation was that the training program itself contained lots of reflection so that the reflection with the diary could go beyond. However, a positive linear trend for reflected experience could be shown in the diary data. Thus, the diary at least seems to be a better way to measure reflected experience than the questionnaire items, because it proved to be prone to changes.

Conclusion.

In sum, the construct of teachers' diagnostic competence has been broadened in this thesis by the results of an empirically tested and validated process model that contains important teacher actions and knowledge for diagnosing students' learning behavior appropriately and for providing didactic action afterwards. A suitable instrument has been developed to measure that competence. Furthermore, important factors for the development of diagnostic competence have been detected. These are a teacher's knowledge about diagnosing learning behavior and a teacher's professional self-concept. On that basis, an amelioration of teacher education and further education concerning the new concept of diagnostic competence has been indicated. In particular, training programs for teachers who are already on the job and did not profit from the newly modified second phase of teacher education have been needed. A first training program has been developed and showed to be effective. Furthermore, standardized diaries have proven to be a suitable way of measuring the course of application of the trained variables at school.

Findings have educational implications for both the assessment of competences and teachers' diagnostic action in the classroom. The case scenario can be applied as an economic, close-to-action instrument to measure teachers' diagnostic competence concerning learning behavior, and the tested model can be used as a code of practice for teachers' diagnostic action in the classroom. There is still much potential in teacher education and further education in the field of diagnosing especially when it comes to learning behavior. Fostering this competence is important for teachers' practical work because diagnosing learning behavior is a main task at school. Combined with expertise research, it will be helpful to foster teachers' professional self-concept as a relevant variable for their competence development in addition to fostering the process of diagnosing learning behavior at different stages of their careers. A modification of teacher education and training programs for teachers

on the job can finally lead to a facilitation and amelioration of teachers' diagnosing action at school contributing to fostering their students learning behavior. Thus, the aim should be to multiply the knowledge by keeping on training teachers and first and foremost by implementing the concept in teacher education. When doing that, the substantial correlation between diagnostic competence and counseling competence should explicitly be considered.

Future prospect.

In further studies, a more representative sample e.g. within a compulsory investigation would be a great advance. Furthermore, the model and scenario test should be validated with objective data from the classroom. Scenario test data could be compared to classroom observations. In particular, postactional variables could be validated with observational data from the classroom that show whether teachers really change their class strategies by teaching particular learning strategies or by observing counseling talks or by analyzing their plans to foster individual students. Additionally, it would be interesting to investigate the correlation between data gathered by the scenario test about diagnostic competence concerning learning strategies and data on teachers' accuracy of diagnosis. Future studies should aim to conduct such a validation in the field.

Besides, future studies should investigate to a greater extent teachers' competence development in a longitudinal design to obtain more information about the real development of competences and different levels of expertise in different stages of a teacher's career. A longitudinal analysis across teachers' careers based on a representative sample could help us to learn more about teachers' competence development. In the cross-sectional design of manuscript 1, it was not possible to investigate that.

Future studies should also aim to generalize the model, instruments, training program and diary for different target groups, because only grammar school teachers were

investigated in the manuscripts contained in this thesis. Thus, further studies should incorporate other school types such as primary schools.

In further training studies, the postaction content should be broadened, e.g. in the course of a modularization of the training program with one module for each phase, whereas the modules, especially the postaction one, should consist of two or more sessions. With more time and more practice with the postaction content, effects in pre-post comparison should occur for postaction variables.

Additionally, in further training studies, a follow-up test that complements the pretest-post-test comparison should be implemented. It could verify the lasting effects indicated in the diary data with the other instruments.

Concerning the diary, future studies should investigate if a longer diary with more open questions to reflect on would lead to an additional intervention effect or if the work on a diary without participating in a training program would show an intervention effect. Furthermore, some postaction diary variables should be reformulated in a further diary study in a way that the measured behavior can in fact be shown every day. Interrupted time series analyses should show significant intervention effects then.

Future studies on training programs and diaries should take the correlation between diagnostic competence and counseling competence into account and foster both competences with regard to their connection. Additionally, the character of the correlation between diagnostic competence and counseling competence should be further investigated. Experimental studies could aim to test whether diagnostic competence is causal for counseling competence.

References

Abs, H.J. (2007). Überlegungen zur Modellierung diagnostischer Kompetenz bei Lehrerinnen und Lehrern. [Considerations to modelling diagnostic competence of teachers]. In M. Lüders, & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* [*Research of teacher education. Competence development and programme evaluation*], (S. 63 - 84). Münster: Waxmann.

Aich, G. (2006). *Kompetente Lehrer. Ein Konzept zur Verbesserung der Konflikt- und Kommunikationsfähigkeit.* Hohengehren: Schneider.

Arnold, K.-H. (1999). Diagnostische Kompetenz erwerben [Acquiring diagnostic competence]. *Pädagogik, 7-8,* 73-77.

Bakkenes, I., Vermunt, J.D., & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction, 20,* 533-548.

Bates, C., & Nettelbeck, T. (2001). Primary school teachers' judgments of reading achievement. *Educational Psychology, 21,* 177–187.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9,* 469-520.

Berliner, D.C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35,* 463–482.

Brante, G. (2009). Multitasking and synchronous work: complexities in teacher work. *Teaching and Teacher Education, 25 (3),* 430–436.

Bruder, S. (2011). *Lernberatung in der Schule. Ein zentraler Bereich professionellen Lehrerhandelns.* Veröffentlichte Dissertation. Technische Universität Darmstadt. Zugriff am 15.03.2011. Verfügbar unter http://tuprints.ulb.tu-darmstadt.de/2432/1/Dissertation_Bruder_Lernberatung_070311.pdf.

Bruder, S., Klug, J., Hertel, S., Kelava, A., & Schmitz, B. (submitted). Counselling Parents in Learning Strategies – Modelling an Important Aspect of Teachers´ Professional Competences.

Butler, D. L., & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-281.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18,* 947-967.

Coladarci, T. (1986). Accuracy of Teacher Judgments of Student Responses to Standardized Test Items. *Journal of Educational Psychology, 78 (2),* 141-146.

Cox, D.D. (2005). Evidence based interventions using home school collaboration. *School Psychology Quarterly, 20* (4), 473-497.

Demaray, M. K., & Elliot, S. N. (1998). Teachers' judgments of students' academic functioning: A comparison of actual and predicted performance. *School Psychology Quarterly, 13,* 8 –24.

Edelenbos, P., & Kubenek-German, A. (2004). Teacher assessment: the concept of `diagnostic competence`. *Language Testing, 21*, 3, 259-283.

Epstein, R.M., & Hundert, E. M. (2002). Defining and assessing professional competence. *American Medical Association, 287 (2),* 226-237.

Epstein, J.L., & van Voorhis, F. L. (2001). More than minutes: Teachers´ roles in designing homework. *Educational Psychologist, 36,* 181-193.

Feinberg, A. B. & Shapiro, E. S. (2003). Accuracy of teacher judgements in predicting oral reading fluency. *School Psychology Quarterly, 18,* 52-65.

Fend, H. (2006). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen [New theory of school. Introduction to the understanding of educational systems]*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Fiedler, K., Walther, E., Freytag, P., & Plessner, H. (2002). *Organizational Behavior and Human Decision Processes, 88 (1),* 527–561.

Grant, A. M., Franklin, J., & Langford, P. (2002). The Self-Reflection and Insight Scale: A New measure of private self-consciousness. *Social Behavior and Personality, 30,* 821-836.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77 (1),* 81–112.

Hecht, S.A., & Greenfield, D.B. (2002). Explaining the Predictive Accuracy of Teacher Judgments of Their Students' Reading Achievement: The Role of Gender, Classroom Behavior, and Emergent Literacy Skills in a Longitudinal Sample of Children Exposed to Poverty. *Reading and Writing: An Interdisciplinary Journal, 15 (7-8),* 789-809.

Helmke, A., Hosenfeld, I., & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften [Comparative tests as an instrument for the improvement of diagnosic competence of teachers]. In R. Arnold, & C. Griese (Hrsg.), *Schulmanagement und Schulentwicklung [Schoolmanagement and school development]* (S. 119-144). Hohengehren: Schneider.

Helmke, A., & Schrader, F.W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching & Teacher Education, 3 (2),* 91-98.

Hertel, S. (2009). *Beratungskompetenz von Lehrern. Kompetenzdiagnostik, Kompetenzförderung und Kompetenzmodellierung*. Münster: Waxmann.

Hitzinger, H. (1987). *Soziale Kompetenz des Lehrers in der Elternarbeit. Analyse und Bewertung eines problemorientierten gruppendynamischen Lehrertrainings zur Vermittlung sozialer Kompetenz in der Lehrerfortbildung.* Frankfurt: Haag & Herchen.

Horstkemper, M. (2004). Diagnosekompetenz als Teil pädagogischer Professionalität [Diagnosis competence as a part of educational professionalism]. In *Neue Sammlung* 2/2004, 201-214.

Jäger, R.S. (2007). *Beobachten, bewerten, fördern. Lehrbuch für die Aus-, Fort- und Weiterbildung [Observe, value, foster. Textbook for education and further education]*. Landau: Empirische Pädagogik.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al., (2003). *Expertise zur Entwicklung nationaler Bildungsstandards. [Expertise for the development of national educational standards]*. Berlin: Bundesministerium für Bildung und Forschung.

KMK (2000). *Rahmenvorgaben für die Einführung von Leistungspunktsystemen und die Modularisierung von Studiengängen.* Beschluss der Kultusministerkonferenz vom 15.09.2000.

Koeppen, K., Hartig, J., Klieme, E., & Leutner, D. (2008). Current Issues in Competence Modeling and Assessment. *Journal of Psychology, 216 (2),* 61-73.

Kretschmann, R. (2009). *Von der Diagnose zum Förderplan – Diagnostik in pädagogischen Handlungsfeldern [From diagnosing to the planning of fostering - Diagnostics in educational spheres of activity]*. Vortrag anlässlich des Symposiums Beobachten-Erkennen-Verstehen-Begleiten, Graz: Pädagogische Hochschule. Zguriff am 03.03.2009. Verfügbar unter http://home.arcor.de/rudolf.kretschmann/Aufs/Aufs%20Uebersichten/Paedagogische%20Diagnostik.htm

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review (28),* 49–57.

Kultusministerkonferenz (KMK). (2004). *Standards für die Lehrerbildung: Bildungswissenschaften.* Beschluss der Kultusministerkonferenz vom 16.12.2004. Zugriff am 15.03.2011. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf

Kunter, M., Klusmann, U., Dubberke, T., Baumert, J., Blum, W., Brunner, M., Jordan, A., Krauss, S., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2007). Linking aspects of teacher competence to their instruction: Results from the COACTIV project. In M. Prenzel (Ed.), *Studies on the educational quality of schools: The final report on the DFG Priority Programme* (pp. 39-59). Münster: Waxmann.

Lee, K., Chiu, S.N., Hasselt, van, C.A. & Tong, M. (2009). The Accuracy of Parent and Teacher Reports in Assessing the Vocabulary Knowledge of Chinese Children with Hearing Impairment. *Language, Speech, and Hearing Services in Schools, 40,* 31-45.

Landesinstitut für Schule und Weiterbildung. (1998). *Fachgutachten. Beratung in der Schule und im Schulsystem. Ergebnisse einer Überprüfung und Anregungen zur weiteren Entwicklung.* Bönen: Verlag für Schule und Weiterbildung.

Lüdtke, O., Köller, O., Marsh, H.W., & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology, 30,* 263–285.

Maclellan, E. (2004). Initial knowledge states about assessment: novice teachers' conceptualisations. *Teaching and Teacher Education, 20,* 523-535.

Marcos, J. J. M., Miguel, E.S., & Tillema, H. (2009). Teacher reflection on action: what is said (in research) and what is done (in teaching). *Reflective Practice, 10 (2),* 191-204.

Marsh, H.W. & Hau, K.-T., (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology, 32,* 151-170.

McElvany, N., Schroeder, S., Richter, T., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H. & Ullrich, M. (in press). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern [Diagnostic skills of teachers concerning the evaluation of pupil's achievements and task difficulty using learning media with instructional images]. *Zeitschrift für Pädagogische Psychologie.*

McGinn, T., Jervis, R., Wisnivesky, J., Keitz, S., & Wyer, P. C. (2008). Tips for Teachers of Evidence-based Medicine: Clinical Prediction Rules (CPRs) and Estimating Pretest Probability. *Journal of General Internal Medicine, 23(8),* 1261–1268.

McLeod, J. (2003). *An introduction to counselling.* Buckingham: Open University Press.

Meisinger, E.B., Bradley, B.A., Schwanenflugel, P.J., & Kuhn, M.R. (2010). Teachers' Perceptions of Word Callers and Related Literacy Concepts. *School Psychology Review, 39 (1),* 54 – 68.

Opdenakker, M.C., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and Teacher Education, 22(1),* 1–21.

Ophuysen, S.v. (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahnempfehlung [Comparison of diagnostic decisions of novices and experts at the example of school career recommendation]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 38 (4),* 154-161.

Perels, F., Dignath, C., & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education, XXIV*, 17-32.

Perry, N.E., Hutchinson, L., & Thauberger, C. (2008). Talking about teaching self-regulated learning: Scaffolding student teachers' development and use of practices that promote self-regulated learning. *International Journal of Educational Research, 47,* 97–108.

Perry, N.E., VandeKamp, K.O., Mercer, L. K., & Nordby, C.J. (2002). Investigating teacher-student interactions that foster self-regulated learning. *Educational Psychologist, 37(1),* 5-15.

Pintrich, P.R., & van de Groot E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82 (1),* 33-50.

Schmitz, B., & Perels, F. (2011). Self-monitoring of self-regulation during math homework behaviour using standardized diaries. *Metacognition and Learning*, DOI: 10.1007/s11409-011-9076-6.

Schmitz, B., & Wiese, B. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31*, 64-96.

Schnebel, S. (2007). *Professionell beraten. Beratungskompetenz in der Schule*. Weinheim: Beltz.

Shulman, L.S., & Shulman, J.H. (2004). How and what teachers learn: a shifting perspective. *Journal of Curriculum Studies, 36,* 257-271.

Sowa, P.A. (2009). Understanding our learners and developing reflective practice: Conducting action research with English language learners. *Teaching and teacher education, 25,* 1026-1032.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz [Accuracy of the judgment of pupil's

characteristics by teachers and the construct of diagnostic competence]. *Zeitschrift für Pädagogische Psychologie,* 19 (1/2), 85-95.

Strasser, J., & Gruber, H. (2003). Kompetenzerwerb in der Beratung: Eine kritische Analyse des Forschungsstands [Competence acquisition in consultation: A critical analysis of the state of research]. *Psychologie in Erziehung und Unterricht, 50,* 381-399.

Trapnell, P. D., & Campbell, J. D. (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology, 76,* 284-304.

Vogt, F., & Rogalla, M. (2009). Developing Adaptive Teaching Competency through coaching. *Teaching and Teacher Education,* 25, 1051–1060.

Wang, M.C. (1973). The accuracy of teacher's predictions on children's learning performance. *The Journal of Educational Research, 66 (10).*

Webber, J., Scheuermann, B., McCall, C., & Coleman, M. (1993). Research on Self-Monitoring as Behavior Management Technique in Special Education Classrooms: A Descriptive Review. *Remedial and Social Education, 14,* 38-56.

Weinert, F.E. (2001). Concept of Competence: A Conceptual Clarification. In D. Rychen & L. Salganik (Eds.). *Defining and Selecting Key Competencies* (p. 45-65). Seattle: Hogrefe and Huber.

Weinert, F. E. (1999). *Concepts of competence.* Contribution within the OECD project Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo). München: Max Planck Institute for Psychological Research.

Wild, E. (2003). Lernen lernen. Wege einer Förderung der Bereitschaft und Fähigkeit zu selbstreguliertem Lernen. *Unterrichtswissenschaft, 31,* 2-5.

Wilson, E. B. (1952). *An Introduction to Scientific Research.* New York: McGraw-Hill.

Winter, F. (2006). Diagnosen im Dienst des Lernens [Diagnoses in the service of learning]. In G. Becker, M. Horstkemper, E. Risse, L. Stäudel, R. Werning & F. Winter, *Friedrich Jahresheft XXIV*. Seelze: Friedrich Verlag.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-35). San Diego, California: Academic Press.

Part 2: Originalia

Article 1

Diagnostic Competence of Teachers:

A Test of a Process Model that Accounts for Diagnosing Learning Behavior

Julia Klug[1], Simone Bruder[2], Augustin Kelava[1], and Bernhard Schmitz[1]

[1]Technische Universität Darmstadt, Darmstadt, Germany

[2]Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt, Germany

Under Review in *Contemporary Educational Psychology* (IF 2009: 1.057)

Author Note

The preparation of this paper was supported by grants SEHM 1538/5-2 from the German Research Foundation (DFG) in the Priority Program "Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes" (SPP 1293).

Please address correspondence concerning this article to: Julia Klug, TU Darmstadt, FB03, Institut für Psychologie, AG Pädagogische Psychologie, Alexanderstraße 10, 64283 Darmstadt, Germany. Phone: +496151/16-6932. Email: klug@psychologie.tu-darmstadt.de

Abstract

Diagnosing the abilities of students is one of the most central tasks that teachers need to perform in order to create an effective class that will meet the needs of their students. So far, accuracy in teachers' judgments has been measured by correlating their judgments with the results of standardized tests. However, an ongoing request exists to shift the focus from diagnosing students' achievements to diagnosing learning behavior in order to allow for didactic action afterwards. The purpose of this study was to develop and test a model of teachers' diagnostic competence that accounts for learning behavior. In a sample of N = 293 participants (93 grammar school teachers, 107 teacher students in their second phase of teacher education, and 93 teacher students in their first phase of teacher education), a three-dimensional model that describes the process of diagnosing learning behavior was tested by means of confirmatory factor analysis, compared to one- and two-dimensional models, and validated by the participants' actual diagnoses in a case scenario. Using multiple regression analysis, knowledge, professional self-concept, and reflected experience were tested as potential predictors for that competence. Additionally, the level of competence in the three subgroups was compared using ANOVA. Results indicate that the three-dimensional process model indeed provides a substantially better fit than the other models, and it is possible to predict an appropriate diagnosis by the model content. Knowledge and professional self-concept were substantial predictors, but reflected experience unexpectedly was not. Teachers and students in the first phase of education differed significantly in levels of competence, and students in the second phase differed from students in the first phase. However, there was no significant difference between teachers and students in the second phase. Future research should aim to foster this competence at different stages of a teacher's career by developing training programs based on the model.

Keywords: diagnostic competence, learning behavior, process, model, teacher education

Diagnostic Competence of Teachers:
A Test of a Process Model that Accounts for Diagnosing Learning Behavior

Diagnosing students' achievements and learning behaviors is one of the most central tasks teachers need to accomplish in addition to having professional knowledge, giving learner-centered instruction, managing their classrooms, interacting with students, and being motivated role models (Kukla-Acevedo, 2009; Opdenakker & Van Damme, 2006). Consequently, teachers are faced with multitasking and managing highly complex work in their daily job routines (Brante, 2009). Applying Weinert's (2001) concept of key competences to the school setting, diagnostic competence is one of those key competences that teachers require. Teachers' diagnostic competence is of importance because "teachers are challenged to meet diverse learning needs and to adapt their teaching to heterogeneous academic ability as well as to multiple interests and motivations" (Vogt & Rogalla, 2009, p. 1051) in order to create an effective class that meets the needs of their individual students.

With regard to their research topic of language teaching, Edelenbos and Kubanek-German (2004) defined teachers' diagnostic competence as "the ability to interpret students' foreign language growth, to skillfully deal with assessment material and to provide students with appropriate help" (p. 260). This definition can be generalized for diagnosing learning behavior cross-curricularly as the ability to interpret students' academic growth and their growth in using learning strategies.

In looking at prior empirical research, it becomes obvious that investigations have focused primarily on students' academic achievements. From the 1970s until today, teachers' diagnostic competence has been operationalized as their ability to accurately judge their students' achievements or task difficulties. So far, accuracy in teachers' judgments has been measured by correlating teachers' judgments with the results of standardized tests (e.g., Bates & Nettelbeck, 2001; Coladarci, 1986; Demaray & Elliot, 1998; Feinberg & Shapiro, 2003;

Hecht & Greenfield, 2002; Helmke & Schrader, 1987; Lee, Chiu, van Hasselt & Tong, 2009; McElvany et. al., in press; Meisinger, Bradley, Schwanenflugel & Kuhn, 2010; Spinath, 2005; Wang, 1973). Thus, in the common approach to assessing teachers' diagnostic competence, the diagnosis of student achievement has been the focus. However, there is an ongoing request in the theoretical literature on teachers' diagnostic competence to shift the focus from diagnosing students' achievements to diagnosing learning behavior in order to allow for didactic action afterwards (Abs, 2007). Diagnoses should not end in themselves, but rather, the diagnoses should enable teachers to foster their students individually and to adapt their classes to the needs of their students (e.g., Fend, 2006; Horstkemper, 2004; Kretschmann, 2009; Vogt & Rogalla, 2009; Winter, 2006). Even if accuracy in judgments is still a key component of teachers' diagnostic actions, it should not be the sole representation of diagnostic competence anymore.

Therefore, the purpose of this study was to develop and test a model of teachers' diagnostic competence that closes the gap between previous empirical research on diagnostic competence and recent theoretical demands.

First, we identified theoretical components of teachers' diagnostic competence by summarizing multiple demands mentioned in the literature; thus building a process consisting of three postulated dimensions. Second, we identified potential predictors of teachers' diagnostic competence that influence competence development. Third and most importantly, we tested the robustness of the claimed three-dimensional model using confirmatory factor analysis and compared this model with a g-factor model and a two-dimensional model. Fourth, we investigated separately the influence of postulated predictors of competence development for (a) a group of teachers, (b) teacher students in the first phase of German teacher education, and (c) teacher students in the second phase of teacher education.[1] Finally,

we tested teachers, student teachers in their second phase, and student teachers in their first phase of teacher education for differences in their levels of diagnostic competence.

1.1. Identifying theoretical components of teachers' diagnostic competence concerning learning behavior

Below, we describe our postulated model of teachers' diagnostic competence that accounts for the domain of diagnosing learning behavior cross-curricularly followed by the potential predictors we assumed for this competence. Our assumptions were based on expertise research.

Koeppen et al. (2008) defined competences as context-specific cognitive dispositions that are acquired and needed to successfully cope with certain task situations in specific domains. They asked for theoretically sound and empirically tested competence models on which valid measures of competence need to be based: "These models have to (a) represent the internal structure of competencies in terms of specific basic skills and abilities, (b) describe different levels of competencies with reference to domain-specific performance, and (c) take into account changes occurring in learning and developmental processes" (Koeppen, 2008, p. 62). In addition to the theoretical framework of the model, the empirical testing should lead to a "methodological-substantive synergy" (Marsh & Hau, 2007).

In our model of teachers' diagnostic competence, we implemented the assessment of these competence characteristics as follows: The domain consisted of teachers' diagnoses. Concerning context specificity, we focused in particular on teachers' diagnoses of pupils' learning behavior both at school and at home. Focusing on learning behavior, we addressed questions such as how pupils deal with tasks; how they do their homework; how they learn at home; which learning strategies they are able to apply, for example, of the ones that are investigated with the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia & McKeachie, 1991); and above all, how much of a self-regulated learner they

are. Teachers have to diagnose their students' abilities to apply these learning strategies. Furthermore, issues of self-regulated learning can be fostered by the teacher when the teacher diagnoses problems with the students (e.g., Perry, Hutchinson, & Thauberger, 2008; Perry, VandeKamp, Mercer, & Nordby, 2002).

In order to specify such a model of teachers' diagnostic competence concerning learning behavior, we identified theoretical components that represent the internal structure of the competence. Jäger (2007) conceptualized diagnosing as a process. Following the terminology used in models of self-regulation (Schmitz & Wiese, 2006; Zimmerman, 2000), we postulated three cyclical dimensions in the diagnosis process; these dimensions take place in a preactional, an actional, and a postactional phase. Each phase represents one dimension of the model. Figure 1 illustrates the process model of the diagnosis of learning behavior.

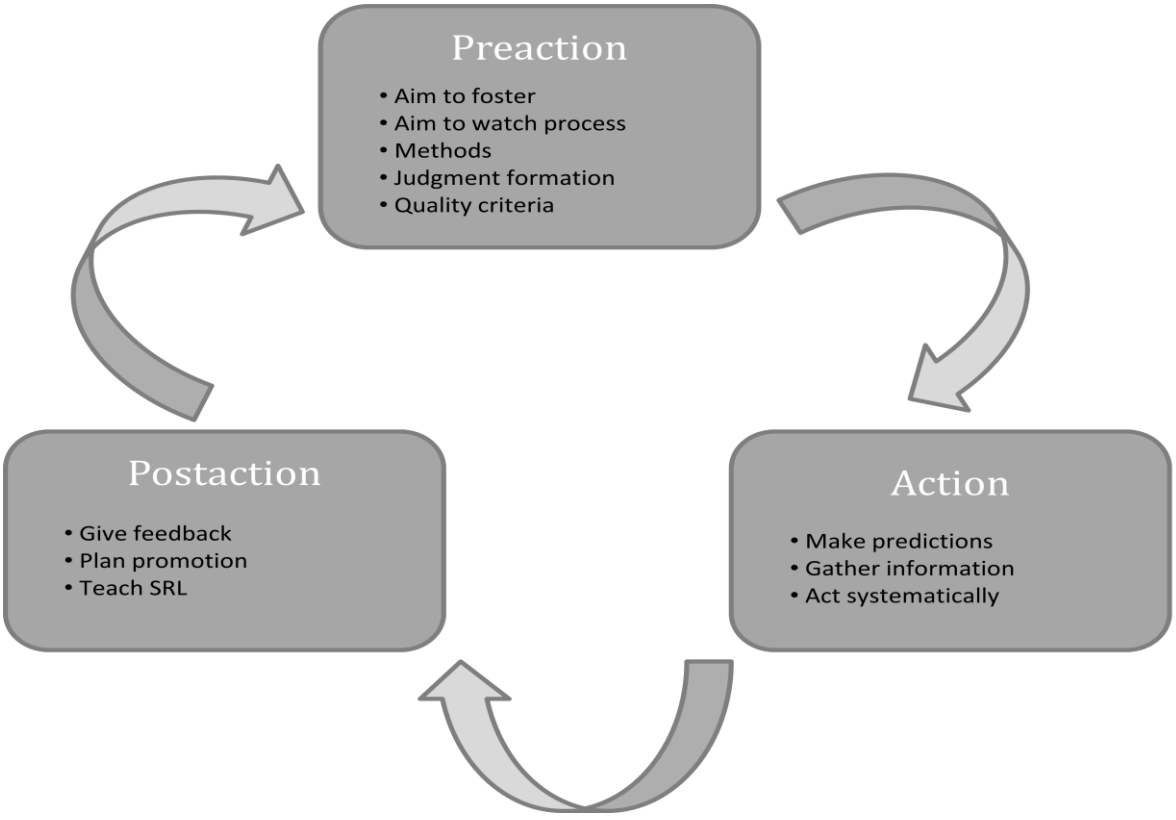


Figure 1. Process model of teachers' diagnostic competence concerning pupils' learning behavior.

The first dimension consists of the preactional phase. In that phase, every diagnosis action—before summing the information to get an actual diagnosis of a pupils' learning behavior—matters. It is especially important that the teacher sets the aim of the diagnosis, which means that the teacher should intentionally aim to watch the individual student's learning process and to foster the student based on the diagnosis (Abs, 2007; Horstkemper, 2004; Kretschmann, 2009). Looking at the student's individual learning process is especially important to develop an individual frame of reference, and with that, to reduce the big-fish-little-pond effect (Lütdke, Köller, Marsh, & Trautwein, 2005). Furthermore, basic diagnostic skills that the teacher possesses are activated in the preactional phase (Strasser & Gruber, 2003). These basic diagnostic skills are composed of knowledge about methods for gathering information about the student (Arnold, 1999; Helmke, Hosenfeld, & Schrader, 2004), knowledge about the psychological quality criteria of tests, and knowledge about judgment formation (Ophuysen, 2006). Regarding methods for gathering information, the teacher should at best not only be familiar with, but should also know how to deal with these methods, and should know the situations in which each method is the most effective. Concerning quality criteria, Maclellan (2004) stated that "it is not sufficient that teachers be procedurally skilled in implementing assessment (necessary as such skills are) but they must also be able to reason about their assessment practice" (p. 524). Her study showed that teachers "do not connect issues of replicability and generalisability (and therefore of the issues of reliability and validity) with assessment method" (p. 530). Knowledge about judgment formation is a construct taken from social psychology where systematic biases due to judgment heuristics influence a person's decision making. Also, within classrooms, such biases can skew teachers' judgments. Fiedler, Walther, Freytag, and Plessner (2002) investigated systematic biases in teachers' judgments of student achievement in a simulated

classroom. If a teacher wants to come to an adequate diagnosis, it is necessary to know theses biases and be aware of them in order to be able to avoid them.

The second dimension of the model consists of the actional phase, in which the actual diagnostic action takes place. Most important in that phase is acting systematically to make a reliable diagnosis. Making a prediction about a student's development and possible underlying learning difficulties seems important. To make a prediction, the teacher has to gather information from different sources and choose the relevant information. Finally, the teacher can interpret the data and come to a concluding diagnosis. He can compare real developments with the ones he predicted in order to make potential changes to his modus operandi for his next diagnosis if he notices differences between his prediction and what really happened. The systematic approach is based upon proceeding scientifically in doing quantitative research (Wilson, 1952). The step of making predictions—in particular, while diagnosing—is adopted from medicine where clinicians can use clinical prediction rules "to predict the most likely diagnosis, prognosis, or response to treatment in a patient based on individual characteristics" (McGinn, Jervis, Wisnivesky, Keitz, & Wyer, 2008, p. 1261).

The third dimension of the model consists of the postactional phase, which begins right after a diagnosis has been made. In this phase, a pedagogical action that follows from the diagnosis should be implemented (Abs, 2007; Kretschmann, 2009). First, giving feedback to the student and the student's parents is a key component of the pedagogical action afterwards—as Hattie and Timperley (2007) concluded, "feedback [to students] is among the most critical influences on student learning" (p. 102). In their conceptual analysis of feedback, the authors gave advice about the conditions under which feedback is the most effective. Likewise, Butler and Winne (1995) stated that "feedback is inherent in and a prime determiner of processes that constitute self-regulated learning" (p. 245). Concerning feedback to parents, making a diagnosis builds the basis for the adequate counseling of parents (Klug,

Bruder, Keller, & Schmitz, in press). Diagnosing also turns out to be one of four dimensions of counseling competence in the counseling model developed by Bruder, Klug, Hertel, Kelava, and Schmitz (submitted) aside from counseling skills, cooperation/perspective-taking, and coping. Second, writing down plans for the individual student's promotion[2] is another content area of the pedagogical action that should be implemented after the diagnosis. In these plans, teachers have to write down (a) the characteristics of the student that should be fostered, for example, motivation, concentration, creativity, ability to think abstractly, and so forth; (b) the student's actual skill level; (c) the goals to be reached; and (d) the measures that will be used to reach these goals. Third, adapting the class as a reaction to the diagnosis by means of teaching appropriate learning strategies and self-regulated learning (SRL) is an important pedagogical action that follows the diagnosis. Several studies have emphasized the relevance of self-regulated learning for students' academic achievement. Pintrich and van de Groot (1990), for example, found that "self-regulation was the best predictor of academic performance […,] which suggests that the use of self-regulating strategies, such as comprehension monitoring, goal setting, planning, effort management and persistence, is essential for academic performance on different types of actual classroom tasks" (p. 38). A study also showed that "it is possible to support self-regulation competencies and mathematical achievement by self-regulation intervention within regular mathematics lessons" (p. 17) held by teachers (Perels, Dignath, & Schmitz, 2009).

Because of the assumed cyclical nature of the model, the three dimensions can influence each other, and in particular, there is a connection between the postactional phase in one diagnosis situation and the preactional phase in a consecutive diagnosis situation. We expected this three-dimensional model with intercorrelations between the phases to fit better than a one-dimensional model of diagnostic competence.

1.2.Potential Predictors of Diagnostic Competence and the Development of Expertise

The following section discusses variables that predict changes in the developmental process and lead to higher levels of administering diagnoses of learning behavior.

What are the relevant variables that lead to a high level of expertise in the diagnosis of learning behavior? In the field of medical diagnosing, Epstein and Hundert (2002) state that competence builds on a foundation of basic clinical skills, scientific knowledge, and moral development. A competent clinician needs to acquire and use knowledge and to integrate data in clinical reasoning to solve real-life problems. Furthermore, he needs the willingness, patience, and emotional awareness to use these skills humanely. Bakkenes, Vermunt, and Wubbles (2010) recently defined teacher learning as an active process in which teachers engage in activities that lead to a change in knowledge and beliefs and/or teaching practices. Strasser and Gruber (2003) state that expertise in the field of counseling is build on knowledge, reflected experience, and personal resources. The so-called COACTIV model developed by Kunter et al. (2007) illustrates teachers' knowledge and beliefs as two key components of teachers' expertise. Shulman and Shulman (2004) integrated also personal variables such as vision and motivation and cognitive variables such as understanding, practice, and reflection into their model of teaching. In particular, reflection on experienced action, which is addressed as an additional deliberate practice, seems to be a key factor for competence development (e.g., Berliner, 2001). The central idea is that "through reflection the teacher better understands and extends his/her professional activity, and that reflecting on teaching problems will lead to new insights for practice" (Marcus, Miguell, & Tillema, 2009, p. 191). It allows practitioners to examine their own clinical reasoning strategies (Epstein & Hundert, 2002). In Clarke and Hollingsworth's (2002) model of teachers' professional growth, the symbiosis of reflection and enactment is the central mechanism for gaining

professional growth, and Sowa (2009) was able to show that reflection was a crucial factor that contributed to changes in teaching.

Based on such expertise and competence development research, we postulated three potential predictors of teachers' diagnostic competence: (a) knowledge about diagnostics, (b) reflected experience in the field of diagnosing, and (c) professional self-concept in diagnosing, which consists of certain attitude variables (motivation to diagnose, attitude toward diagnosing, self-efficacy in diagnosing, and conscientiousness) and serve as personal resources the teacher brings with him. We expected these three variables to be substantial predictors of teachers' diagnostic competence.

1.3.The present study

A large body of research exists on the accuracy of teachers' diagnostic competence, but to the best of our knowledge, there is not yet any empirical research on the process of diagnosing learning behavior and no model of what a diagnostically competent teacher needs to know and to do. The purpose of the present study was to test a newly developed model of teachers' diagnostic competence that accounts for the learning behavior of the student and the process of diagnosing it. We predicted that our claimed three-dimensional model would fit the data better than a one-dimensional model and that the three variables postulated for gathering expertise would predict the values of diagnostic competence. Furthermore, by using the proposed model, we expected that a teacher's diagnostic competence test score would predict whether or not the teacher was able to make an appropriate diagnosis. Finally, the level of competence of teachers with professional experience was compared to the level of teacher students in their first and second phases of German teacher education. We expected teachers to be more competent diagnosticians than students in the second phase, who, in turn were expected to be more competent than students in the first phase.

Method

2.1. Participants

A total of *N* = 293 participants from different German federal states took part in the study (93 grammar school teachers, 107 teacher students in their second phase of teacher education, and 93 teacher students in their first phase of teacher education). One hundred seventy-three participants were female (59%), 110 were male (38%), and 10 did not specify their genders. One hundred seventy-five participants (60%) were between the ages of 20 and 29, 66 were 30 – 39 years of age (23%), 22 were 40 – 49 years of age (8%), 25 were 50 – 69 (9%), and three (1%) were 60 years or older. Two persons did not specify their ages. Multiple answers were possible on subject questions. All requested subjects were represented with a predominance of languages (143) and social sciences (144), followed by natural sciences (86), mathematics (60), and arts/music/sports (50). Seventeen percent had already taken part in a further educational program on diagnostics; 83% had not. With respect to years of experience in the teaching profession, 30% of the subsample of teachers had been school teaching for 0 – 5 years, 26% had been school teaching for 6 – 10 years, 17% for 11 – 15 years, 9% for 16 – 20 years, 2% for 21 – 25 years, another 2% for 26 – 30 years, and 14% had been teaching for more than 31 years. In the subsample of teacher students in the second phase of teacher education, 43% were in their first provisional teaching period and 56% were in the second period. One gave no information. The semester number in the subsample of teacher students in their first phase of teacher education at university ranged from the third to fourteenth semester.

2.2.Materials

Participants completed three tests. Diagnostic competence concerning learning behavior was measured by a scenario test with open questions based on the model.

Knowledge of diagnostics as a predictor was measured by a multiple-choice knowledge test and the two other predictors, professional self-concept and reflected experience, were measured by means of a questionnaire via self-assessment. Some demographic data were recorded additionally from the questionnaire.

1.2.1. Scenario Test

A scenario test was our method of choice to measure diagnostic competence concerning learning behavior based on the model, and was as close to assessing real action as efficiently as possible. The test consists of a case description of a pupil who has certain difficulties in self-regulated learning, leading to a descent in his achievements. The case was constructed based on a matrix of characteristics. This matrix classifies possible characteristics of diagnostic situations with different degrees of difficulty. We chose a medium level of difficulty and integrated the corresponding characteristics into the scenario. The teacher is asked to put himself in the place of this student's teacher. The case description is followed by 12 questions with space for open answers. Eleven of these questions represent the content of the model. They ask for the teacher's behavioral processing when diagnosing the student's learning behavior. There is one question for each area of model content (aiming to foster, aiming to watch individual processes, activating knowledge about methods, activating knowledge about judgment formation, activating knowledge about quality criteria, making predictions, gathering information, acting systematically, giving feedback, planning promotion, teaching learning strategies). A twelfth question, which addresses the final diagnosis the teacher comes to, is included to validate the model components. The idea is that the accuracy of that diagnosis can be predicted by the sum value of the other questions, which are based on the model content. If the values in the model content can predict the accuracy of the final diagnosis, the behavioral processing that the teacher goes through while diagnosing can be seen as a valid predictor of the teacher's diagnostic accuracy. Answers to each

question are rated from 0 – 3 points by three independent raters whose judgments are averaged at the end. The raters were given a handbook with detailed references about how to rate the answers. Inter-rater reliabilities of each question were good with values all between ICC = .67 and ICC = .95. Table 1 summarizes the questions that followed the case scenario.

Table 1

Questions Following the Case Scenario

Instructions: Please answer the following questions about the case scenario. Using the information you were given...
1. You have already received a picture of Marco's working behavior in the lessons by systematically observing Marco at work in groups and in quiet work phases. Do you need further information? If yes, what information? Where do you get this information from?
2. Which goal do you pursue while observing Marco?
3. If you want to assess Marco's achievements, what do you compare his performance level to?
4. From his work in former classes and reports as well as your initial experiences with Marco, he seems to be an intelligent boy. His achievements until recently were good or even very good. Now you find out in your observations that Marco seems sometimes a little bit distracted during lessons. Furthermore, you have repeatedly observed that a certain type of exercise gives Marco problems and he tends to make mistakes. You have already often experienced this with scholars of his age. What are some reasons you would propose for his drop in performance?
5. Do you make sure that your impression that you have formed about Marco is objective? If yes, how?
6. Do you make predictions about Marco's further development? If yes, why?
7. To come to a final judgment about Marco's learning difficulties, you must find out the cause of his problems. How do you proceed with the diagnosis?
8. Which information do you implicate for making your diagnosis?
9. Now you have a precise impression about Marco's difficulties. What is the cause of his learning problems in your mind?
10. Do you give Marco's parents feedback about his learning problems and supposed causes? If yes, what do you report back?
11. What do you do next, after you have fathomed the causes of Marco's learning difficulties?
12. You would like to prevent learning difficulties such as Marco's as much as possible. What do you communicate for this purpose in the lessons?

1.2.2. Knowledge Test

The knowledge test measures basic knowledge of diagnostics as a predictor of diagnostic competence with 11 multiple-choice questions, which are also based on the model content. If multiple answers were possible, it was mentioned after the specific question. Item difficulties were between .33 and .82 and therefore fell within an acceptable range for items that are intercorrelated (e.g., Ramsey & Reynolds, 2000).

1.2.3. Questionnaire

Professional self-concept and reflected experience in diagnostics as postulated predictors of diagnostic competence were measured by a questionnaire. The questionnaire additionally surveyed some demographic data.

1.2.3.1. Professional self-concept

Professional self-concept in diagnostics was measured with 12 items on a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). There were three items that predominantly asked for attitude toward diagnosing: "It is important to me to evaluate my students accurately"; "If I am able to evaluate my students properly, I will be able to create a better class"; and a negatively worded item "To foster a student, it is not necessary to be a good diagnostician." Another three items predominantly asked for motivation to be a good diagnostician: "I am motivated to broaden my diagnostic skills"; "I like to be able to diagnose appropriately"; and "I am motivated to look into reasons for the learning problems of my students."

Another three items of the scale asked for self-efficacy in diagnostics: "I am sure that I can provide something for the individual student due to my diagnosis"; "I know that I am good in diagnosing learning behavior"; and "I am sure to be able to find the reasons for the learning problems of my students." Additional three items of the professional self-concept scale assess teachers' conscientiousness. The items were taken from the BFI-K (Rammstedt & John, 2005): "I complete tasks soundly"; "I am proficient and work briskly"; and "I make plans and accomplish them." The intern consistency of the overall professional self-concept scale was good with $\alpha = .77$.

1.2.3.2. Reflected Experience

Reflected experience in diagnostics was measured with four items, again using a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). The self-constructed items were: "I regularly consider whether I evaluate my students appropriately"; "I am often concerned about whether I appropriately appreciate my students"; "If I diagnose a student with something, I reflect on whether I was correct"; and "I often diagnose in class and try to reconsider my diagnoses afterwards." The reflected experience scale also had a good intern consistency with $\alpha = .75$.

1.2.3.3. Demographic Data

The demographic data section requested participants' age (in 10-year increments), gender, and school subjects (mathematics, natural sciences, languages, social sciences, arts/music/sports). For school subjects, multiple answers were possible. Additionally, participants were asked whether they had already taken part in some kind of further educational program on diagnosing (yes/no). All participants had to answer two items about job experience on a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*): "I am very experienced in the area of pedagogical diagnostics" and "I have already performed many diagnoses in class." The subsample of teachers was asked for their years in the teaching profession (in 5-year increments up to more than 31 years), whereas the subsample of student teachers in the second phase of teacher education were asked whether they were in the first or second provisional teaching period. Both groups were additionally asked on a 6-point Likert scale whether the area of diagnostics was intensively covered in their first and second phases of teacher education. The subsample of teacher students in the first phase of teacher education was asked for their semester. They were asked if the area of diagnostics was intensively covered in their studies at university.

2.3. Procedure

In order to recruit participants, we contacted German grammar schools and comprehensive schools with a college preparatory track as well as teachers' training colleges that are responsible for the second phase of teacher education and universities that educate teacher students. Participation was voluntary, and each participant received a 10 euro voucher to order books. Teachers from schools in the federal state of Hesse were offered the opportunity to receive credit points for taking part in the study. Participants were reassured that their information would be kept confidential and that participation was anonymous. The instruments were available in an online version and in a parallel paper-and-pencil version in order to obtain as many participants as possible. For both variants, there were three slightly different forms concerning the questionnaire section on demographics; one for each subgroup (teachers, teacher students first phase, teacher students second phase). Completing the three instruments took the participants approximately half an hour.

Results

1.1. Descriptive statistics

Means, standard deviations, minima, and maxima were calculated for the questionnaire items on job experience, reflected experience, and professional self-concept, for the total knowledge values on the multiple-choice knowledge test, and for the total score and the dimensions of diagnostic competence in the case scenario. Table 2 summarizes these descriptive statistics.

Table 2

Descriptive Statistics

Item	Instrument	*M*	*SD*	Min	Max	Total possible score
I am very experienced in the area of pedagogical diagnostics	Questionnaire: Job experience	2.73	1.17	1	6	1-6
I have already performed many diagnoses in class	Questionnaire: Job experience	3.30	1.38	1	6	1-6
The area of diagnostics was intensively covered in the second phase of teacher education	Questionnaire: Teacher education	2.98	1.60	1	6	1-6
The area of diagnostics was intensively covered in the first phase of teacher education	Questionnaire: Teacher education	1.90	1.31	1	6	1-6
Reflected experience scale	Questionnaire: Reflected experience	4.30	0.97	1	6	1-6
Professional self-concept scale	Professional self-concept	4.78	0.51	1	6	1-6
Knowledge sum score	Knowledge test	28.42	6.14	8	51	62
Diagnostic competence sum score	Case scenario	14.28	4.59	0	26	36
Preactional dimension	Case scenario	1.50	0.54	0	2.83	3
Actional dimension	Case scenario	1.18	0.57	0	2.67	3
Postactional dimension	Case scenario	1.04	0.62	0	3	3

Concerning the questionnaire, for the items asking whether the area of diagnostics was intensively covered in teacher education, the mean values were rather low, whereas diagnostic action was needed very often in class. The scales addressing experience and professional self-concept were answered with rather high values and a low variance.

Of a total possible score of 62 on the knowledge test, just half of the possible points were reached. Concerning the sum score of diagnostic competence in the case scenario, with $M = 14.28$, the mean value was even lower than half of the total possible score of 36 points. Looking at the descriptive values of the three postulated dimensions, for each dimension, constructed as the mean of the content items, the possible maximum was 3. For the preactional dimension, the mean score was $M = 1.50$ ($SD = 0.54$). The mean values for the

other two dimensions were lower. The mean score of the actional dimension was $M = 1.18$ ($SD = 0.57$) and the mean score of the postactional dimension was $M = 1.04$ ($SD = 0.62$).

1.2. Confirmatory factor analysis

Like Muthén & Muthén (2010) point out that the relationships between a set of observed variables and a set of continuous latent variables can be studied that with the help of confirmatory factor analysis (CFA). To compute confirmatory factor analysis, we used MPlus. Because the Lilliefors test, a special case of the Kolmogorov Smirnov test, with which we assessed whether the data were normally distributed, revealed that the data were not normally distributed, we used maximum likelihood with robust standard errors and chi-square (MLR) as the estimator. That estimator provides Satorra-Bentler-corrected standard error estimates and chi-square values. Correlations among the factors were freely estimated. We expected the factors to correlate with each other because of the anticipated process structure.

The three-dimensional model showed that the model fit the data very well. The chi-square test was not significant, $\chi^2(36) = 47.704$, $p = .092$. Hu and Bentler (1998, 1999) have suggested using the SRMR to assess model fit because of its sensitivity to simple model misspecification. They suggested that target values of the SRMR should be less than .08 in order to indicate adequate model fit. The RMSEA too indicates adequate fit with values of less than .08 (Browne & Cudeck, 1992). With an SRMR of .045 and an RMSEA of .033, our model showed good model fit. The comparative fit index (CFI) developed by Bentler (1990) is an incremental fit index (Hoyle & Panter, 1995; Hu & Bentler, 1999), which is sensitive to misspecified factor loadings. It assesses the improvement of fit of the hypothesized model over the null model and ranges from 0 to 1. Values greater than .95 have recently been advocated (Hu & Bentler, 1999) to indicate improved fit as an increase from earlier target values greater than .90 (Hoyle & Panter, 1995). Our model had a CFI of .954, exceeding the newly advocated cut-off value. Figure 2 shows the resulting measurement model.

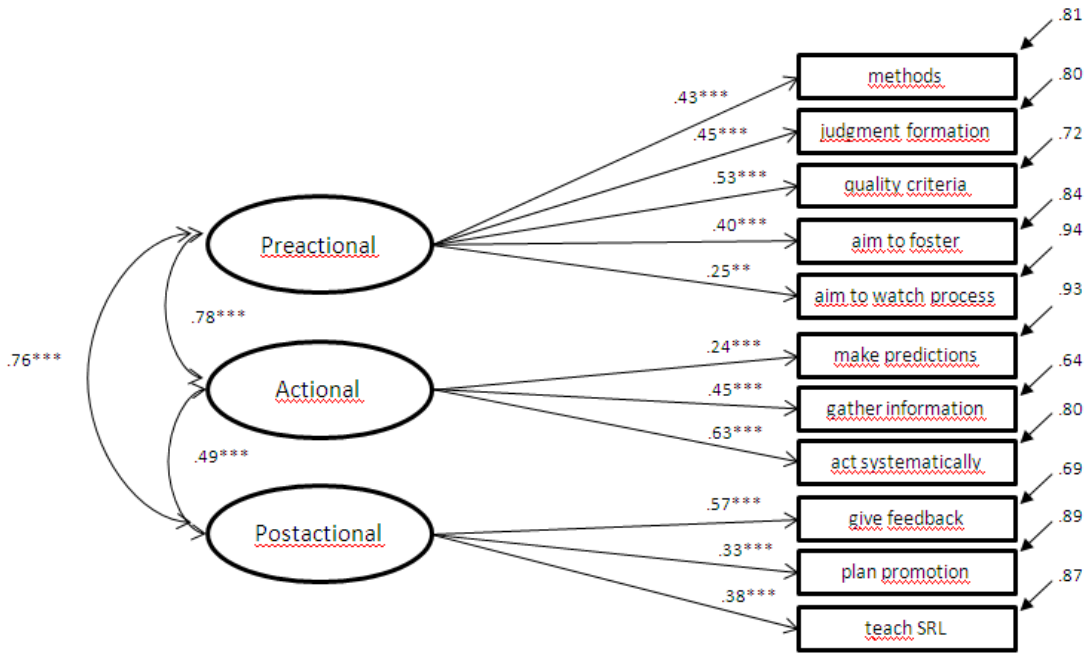


Figure 2. Three-dimensional model of diagnostic competence.

The factor structure confirmed the three dimensions, which represent the phases of the diagnosing process. The dimensions were substantially intercorrelated as expected because of the process structure of diagnosing. The preactional phase consisted of the postulated content variables regarding the aim of the diagnosis and basic diagnostic skills that are activated: activating knowledge about methods, activating knowledge about judgment formation, activating knowledge about quality criteria, aiming to foster, and aiming to watch the individual's process. The actional phase consisted of the postulated content variables that represent processing while actually diagnosing: making predictions, gathering and choosing relevant information, and acting systematically. The postactional phase consisted of the expected content variables that describe pedagogical action afterwards: giving constructive feedback to students and parents, planning to foster the individual student, and teaching adequate learning strategies in class.

1.2.1. Comparison of the one-dimensional and two-dimensional versus three-dimensional models

In addition to the test of the fit of our claimed three-dimensional process model with intercorrelations between the factors, we wanted to test whether a g-factor model and a two-factor model could fit the data as well. Table 3 shows the model fit indices of the models.

Table 3

Model Comparison

	n	χ^2	*p*	*df*	χ^2/df	CFI	RMSEA	SRMR	AIC	BIC[1]
One-dimensional model	291	60.226	.016	39	1.544	.954	.043	.051	7895.501	7914.582
Two-dimensional model	291	60.536	.012	38	1.593	.911	.045	.051	7897.050	7916.633
Three-dimensional model	291	47.704	.092	36	1.325	.917	.033	.045	7887.518	7908.449

[1] Sample-size adjusted BIC

As table 3 shows, the chi-square values of the g-factor and the two-dimensional models were higher than for the three-dimensional model and were significant, which means that the g-factor and the two-dimensional models differed significantly from the empirical data in contrast to the three-dimensional model. The CFI value of the g-factor model did not reach the cut-off value of .95, whereas the CFI of the two-dimensional and three-dimensional models did. Each model met the cut-off criteria of the RMSEA and SRMR of .08, but the values of the three-dimensional model were lower than those of the other two models, indicating a better fit of the three-dimensional model. As further indices for model comparison, we consulted the AIC and BIC. Those indices quantify the degree to which the given model represents an improvement over comparison models (McCoach & Black, 2008).

Following Zucchini (2000), the BIC regards every competing model as the possible 'true' model before estimating the likelihood that the model in question is, indeed, the correct one. For the AIC, the prediction of future data is the key criterion of the adequacy of a model (Kuha, 2004). McCoach and Black (2008) believe that the combined use of the AIC and BIC in conjunction with chi-square difference tests can be quite informative. Thus, we chose to report both indices in addition to a chi-square difference test. Both indices indicate a better fit if they are smaller than the respective index from a competing model. Looking at these indicators, we found that the AIC and BIC of the three-dimensional model were smaller than the ones of the g-factor model and the two-dimensional model, again, indicating a better fit of the three-dimensional model.

To calculate a chi-square difference test, we needed a test that accounts for non-normal data. Muthén and Muthén (2010) describe a chi-square difference test for the Satorra-Bentler-scaled chi-square developed by Albert Satorra (2000). In that test, the usual normal-theory chi-square statistic is divided by a scaling correction to better approximate a chi-square distribution under non-normality (Muthén & Muthén, 2010). The steps that are needed to compute the chi-square difference test that accounts for the used MLR estimator can be found on the MPlus website (Muthén & Muthén, 2011). The calculated empirical value TRd = 11.158 ($df_{\text{diff}} = 3$, $p < .05$) was higher than the critical value. Consequently, the models differed significantly.

By assessing every reported index and the chi-square difference test, we came to the conclusion that the three-dimensional model fit the data significantly better than the g-factor model and the two-dimensional model.

1.3. Validation

In the scenario test, we included a twelfth question, which did not respond to a content variable of the model, but rather asked for the teacher's final diagnosis of the student's

learning behavior. That question was added as a validation question. The idea was that the accuracy of that diagnosis could be predicted by the values of the three dimensions of the model content. Consequently, we expected to be able to predict an appropriate diagnosis by the teacher's diagnostic competence score on the test based on the components of the model. To test that hypothesis, we used a multiple regression analysis with the dimension scores from the case scenario as predictors and the accuracy of the diagnosis measured by that validation question as the criterion. Multiple regression analysis showed that it is possible to significantly predict the correctness of the diagnosis from the teacher's preactional, actional, and postactional dimension scores with $R^2 = .16$. Table 4 shows the results of the multiple regression analysis.

Table 4

Multiple Regression Analysis to Validate the Model

	R^2	Preactional	Actional	Postactional
Accuracy of diagnosis	16%***	.14*	.27***	.14**

*$p < .05$. **$p < .01$. ***$p < .001$.

1.4. Predicting diagnostic competence separated by groups

Based on expertise and competence development research, we expected the following three variables to be substantial predictors of teachers' preactional, actional, and postactional diagnostic competence: (a) knowledge about diagnostics, (b) reflected experience in the field of diagnosing, and (c) professional self-concept in diagnosing. Using multiple regression analyses, we tested their prognoses of the three dimensions of diagnostic competence separately for the three groups that consisted of teachers, student teachers in their second phase, and student teachers in their first phase of German teacher education. Table 5 shows the results of the multiple regression analyses separately for the three groups.

Table 5

Multiple Regression Analyses Separated by Groups

		R^2	Professional self-concept	Knowledge	Reflected experience
Teachers	Preactional	11%	.32*	-	-
	Actional	14%	.27**	-	-
	Postactional	9%	.23*	.21*	-.24*
Teacher students second phase	Preactional	3%	-	-	-
	Actional	11%	-	.25*	-
	Postactional	1%	-	-	-
Teacher students first phase	Preactional	12%	-	.32**	-
	Actional	16%	-	.44**	-
	Postactional	10%	-	.28*	-

*$p < .05$. **$p < .01$. ***$p < .001$.

In the group of teachers, their professional self-concept was the best predictor for all three dimensions of diagnostic competence (preactional: $\beta = .32$, $p < .05$, $R^2=11\%$; actional: $\beta = .27$, $p < .01$, $R^2=14\%$; postactional: $\beta = .23$, $p < .05$; $R^2=9\%$). This means that the more motivated, interested, self-efficient and conscientious the teacher is in diagnosing, the better he is in preactional, actional, and postactional diagnostic competence and vice versa. The teacher's knowledge can also significantly predict his postactional diagnostic competence ($\beta = .21$, $p < .05$). Reflected experience unexpectedly appeared to be a negative predictor of postactional diagnostic competence ($\beta = -.24$, $p < .05$).

In the group of teacher students in their second phase of teacher education, only the actional dimension of diagnostic competence was predicted significantly by knowledge, explaining 11% of the variance ($\beta = .25$, $p < .05$). Reflected experience and professional self-

concept did not add significantly to the prediction. The other dimensions had no significant predictors.

In the group of teacher students in their first phase of teacher education, the only variable that could significantly predict each dimension of diagnostic competence was knowledge, indicating that the higher a student's knowledge in diagnostics, the higher that student's competence was in all three dimensions and vice versa (preactional: $\beta = .32$, $p < .01$, $R^2=12\%$; actional: $\beta = .44$, $p < .01$, $R^2=16\%$; postactional: $\beta = .28$, $p < .05$, $R^2=10\%$).

1.5. Expertise levels of teachers and student teachers in their second and first phases of teacher education

Finally, the level of competence of teachers with professional experience was compared to the levels of teacher students in their first and second phases of German teacher education. We expected teachers to be more competent diagnosticians than students in the second phase who in turn were expected to be more competent than students in the first phase on all three dimensions as well as on the diagnostic competence sum score. To test this, we calculated an ANOVA with group as the independent variable and the three dimensions and sum score of diagnostic competence as dependent variables. Table 6 shows the results of this analysis.

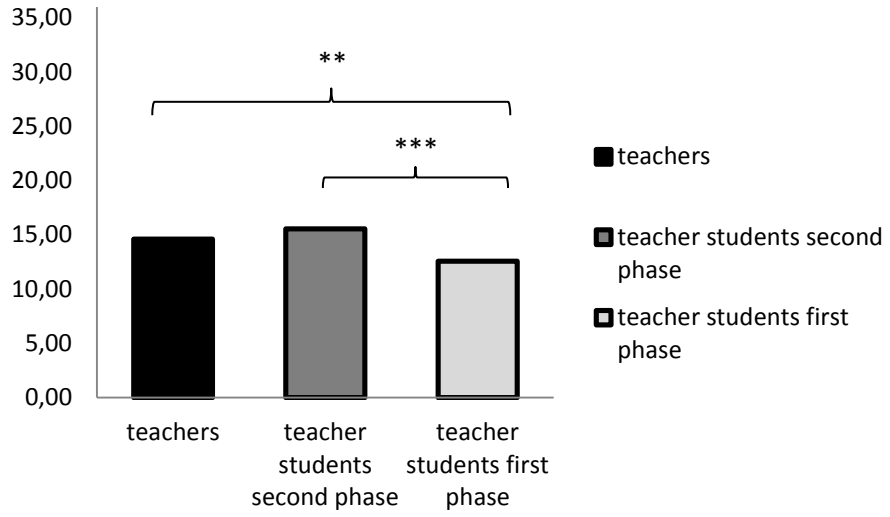
Table 6

ANOVA with Group as the Independent Variable and Sum Score and Dimensions of Diagnostic Competence as Dependent Variables

V	IV	*df*	*F*	Eta^2
Diagnostic competence sum score	Group	2	11.66***	.08
Preactional dimension	Group	2	27.04***	.16
Actional dimension	Group	2	1.18	.01
Postactional dimension	Group	2	7.35**	.05

*$p < .05$. **$p < .01$. ***$p < .001$.

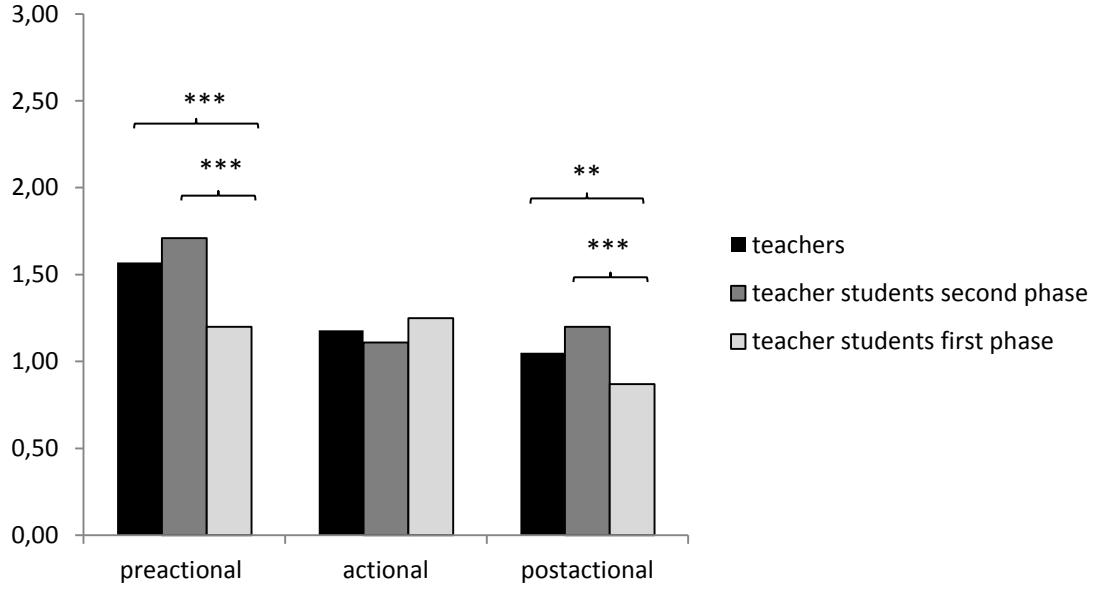
The three groups differed significantly on their diagnostic competence overall score, $F(2) = 11.66$, $p < .001$, $\eta^2 = .08$, as well as on the preactional dimension, $F(2) = 27.04$, $p < .001$, $\eta^2 = .16$, and on the postactional dimension, $F(2) = 7.35$, $p < .01$, $\eta^2 = .05$, but not on the actional dimension. Contrasts further showed the specific differences between the three groups. They are illustrated in Figure 3 for the sum score and in Figure 4 for the dimensions of diagnostic competence.



*$p < .05$. **$p < .01$. ***$p < .001$.

Figure 3. Contrasts between the three groups with the overall score as dependent variable.

As postulated, there was a highly significant difference in the diagnostic competence sum score between teachers and teacher students in the first phase of teacher education and between teacher students in the second phase and teacher students in the first phase. But, unexpectedly, there was no significant difference between teachers and teacher students in the second phase of teacher education. The values of teacher students in the second phase were even a little higher than the ones of experienced teachers. We found the same pattern for the preactional and postactional dimensions.



**p < .05. **p < .01. ***p < .001.*

Figure 4. Contrasts between the three groups with the dimensions as dependent variables.

Discussion

This study tested a newly developed model of teachers' diagnostic competence that accounts for students' learning behavior and the process of diagnosing it. The first research goal was to test our postulated three-dimensional model against one- and two- dimensional models whereby we predicted that the three-dimensional model would fit the empirical data best.

As a second research goal, we expected to be able to predict appropriate diagnoses from teachers' diagnostic competence test scores based on the components of the model.

Our third research goal was to test whether the three variables postulated for gathering expertise would predict the values of diagnostic competence.

Finally, as a fourth research goal, we tested whether teachers with professional experience were more competent diagnosticians than students in the second phase of teacher education who in turn were expected to be more competent than students in the first phase.

Concerning the first research goal, results indicate that the three-dimensional model indeed provides a substantially better fit than a one- or two-dimensional model. Although the one- and two-dimensional models are more economical, the three-dimensional structure fits the empirical data best. We argue that this better fit is due to the multidimensional structure of diagnostic competence concerning learning behavior, even if the three dimensions are each substantially correlated with one another. These correlations in turn emphasize the process character of diagnosing consistent with Jäger (2007), because each dimension has an influence on the other ones. Further studies can take into account the investigation and fostering of the diagnosis of learning behavior based on the model.

As expected, we were able to predict appropriate diagnoses from teachers' diagnostic competence test scores that were based on the model. This leads to the conclusion that teachers who follow the steps of the diagnosis process advised by the model in fact come to a more adequate diagnosis about students' learning behavior and vice versa. This can be seen as a validation of the variables comprised by the model. Nevertheless, a validation of the model and case scenario with objective data from the classroom would be preferable. Case scenario data could be compared to classroom observations. In particular, postactional variables could be validated with observational data from the classroom that show whether teachers really change their class strategies by teaching particular learning strategies or by observing

counseling talks or by analyzing their plans to foster individual students. Additionally, it would be interesting to investigate the correlation between case scenario data about diagnostic competence concerning learning strategies and data on teachers' accuracy of diagnosis. Future studies should aim to conduct such a validation in the field.

Concerning the third research goal, the test of the predictors of diagnostic competence was partly confirmed. Professional self-concept in diagnosing and knowledge about diagnostics proved to be substantial predictors of teachers' diagnostic competence. Professional self-concept in diagnosing is the best predictor followed by knowledge about diagnostics in the group of teachers, but not in the other two groups. This means that the more motivated, interested, self-efficient, and conscientious the teacher is in diagnosing and the more knowledge he has about diagnostics, the better he is in preactional, actional, and postactional diagnostic competence and vice versa. In the other two groups of teacher students in the first and second phases of teacher education, knowledge about diagnostics was the only relevant predictor of their diagnostic competence. The assumption arises that professional self-concept does not become relevant for competence development until the teacher is finally on the job, but then it becomes more important than knowledge for their further development. These findings that cognitive variables such as knowledge about diagnostics and personal variables such as professional self-concept are relevant for the development of competence are in line with findings from expertise research (e.g., Bakkenes, Vermunt, & Wubbles, 2010; Epstein & Hundert, 2002; Shulman & Shulman, 2004; Strasser & Gruber, 2003) and the COACTIV model (Kunter et al., 2007). Unexpectedly, reflected experience appears to be a negative predictor of postactional diagnostic competence in the group of teachers, whereas it is not a significant predictor in the other two groups. This finding does not support research on reflected experience, which has usually been found to be a crucial predictor of competence (Berliner, 2001; Bruder, Klug, Hertel, et al., submitted; Epstein & Hundert, 2002; Clarke &

Hollingsworth, 2002; Marcus, Miguell, & Tillema, 2009; Strasser & Gruber, 2003; Sowa, 2009). Rumination can be considered as a possible explanation for the small negative correlation between reflected experience and competence measures in the group of teachers. This means that teachers who state that they reflect a lot on their diagnostic actions do not come to an action in the diagnosis process and vice versa. A general problem here lies in the measurement of reflected experience by self-reports, which are prone to social desirability. With $M = 4.30$ ($SD = 0.97$), teachers estimated their reflected experience to be rather high with low variability. In further studies, the measurement of reflected experience should be improved rather than considering reflection as not important for competence development. For that reason, we plan to manipulate reflected experience in a training program with standardized diagnosis diaries, which should stimulate reflection so that we can examine the influence of reflected experience on diagnostic competence.

Our fourth hypothesis that teachers with professional experience will be more competent diagnosticians than students in the second phase of teacher education who in turn were expected to be more competent than students in the first phase was confirmed in part. As postulated, there was a highly significant difference between teachers and teacher students in the first phase of teacher education and between teacher students in the second phase and teacher students in the first phase. But, unexpectedly, there was no significant difference between teachers and teacher students in the second phase of teacher education. This finding is in line with results of a study that compared the counseling competence of these groups (Bruder, Klug, Hertel, et. al., submitted). We can exclude that the effect is due to motivation because we measured motivation and tested the groups for differences in their motivation, and there were none. An explanation for the good result of teacher students in their second phase of teacher education can be found in a change in German teacher education a few years ago when the second phase was modularized. There are existing modules now that broach the

issue of diagnosing, fostering, and counseling, and these seem to be very effective. Nevertheless, teachers on the job who did not get that kind of education lack knowledge about the new content and perhaps compensate for that lack of knowledge with a more professional self-concept, which in the end leads to comparable values in our case scenario. Future studies should investigate to a greater extent teachers' competence development in a longitudinal design to obtain more information about the real development of competences and different levels of expertise in different stages of a teacher's career. In the cross-sectional design of our previous study, we were not able to investigate that. Additionally, there is a need to develop and evaluate on-the-job teacher training programs that base the diagnosing of learning behavior on the model, and that foster professional self-concept, knowledge about diagnosing, and reflected experience, even though reflected experience was not a significant predictor in our study. Considering descriptive results, there is also a need for change in the first phase of teacher education, in which the diagnosing of learning behavior has rarely been addressed until now (M = 1.90, SD = 1.31), which can be seen in the descriptive statistics of some questionnaire items (see table 2). Generally, there is a high need for fostering diagnostic competence across all groups. Participants had a mean of M = 14.28 (SD = 4.59) out of 36 possible points with a maximum of 26 points and a minimum of 0 points in the case scenario sum score, so there is much leeway left to arrive at an optimal result.

The instruments we developed and used in this study proved to be appropriate. All instruments showed satisfactory quality criteria such as high inter-rater reliabilities for the scenario test, medium item difficulties for the knowledge test, and acceptable internal consistencies for the questionnaire. The most important instrument, which was used to measure diagnostic competence concerning learning behavior, is both close to measuring real teacher behavior while still being economical. It sure does not measure real action, but, in any case, it approaches the action the teacher would show in a real situation. There is just the

unsolved problem of whether reflected experience can be measured in any way other than self-assessment, a problem that we already mentioned earlier. Despite this, the instruments can be used in further studies in which we will use pre and post measurements to assess whether they are sensitive to changes due to interventions.

When it comes to the procedure, the greatest limitation of this study needs to be discussed. Because of the voluntary participation, the sample is not representative. The assumption stands to reason that the participants are the rather motivated ones out of the population of teachers and maybe even the ones with rather high diagnostic competence. Nevertheless, the values even of this sample were rather low or medium on the scenario test. It is to be assumed that values in the population are even lower than in the investigated sample.

1.1. Educational Implications

Findings have educational implications for both the assessment of competences and teachers' diagnostic action in the classroom. The case scenario can be applied as an economic, close-to-action instrument to measure teachers' diagnostic competence concerning learning behavior, and the tested model can be used as a code of practice for teachers' diagnostic action in the classroom. Furthermore, implications for teacher education and further education are especially relevant. Consistent with requirements for diagnosing at school, in the theoretical literature, the construct of diagnostic competence has been broadened by the addition of the new process model on diagnosing learning behavior. A change in the topic of diagnosing in teacher education can be based on this model. There is still much potential in teacher education and further education in the field of diagnosing especially when it comes to learning behavior. Fostering this competence is important for teachers' practical work because diagnosing learning behavior is a main task at school. Combined with expertise research, it will be helpful to foster teachers' professional self-concept as a relevant variable for their

competence development in addition to fostering the process of diagnosing learning behavior at different stages of their careers. A modification of teacher education and training programs for teachers on the job can finally lead to a facilitation and amelioration of teachers' diagnosing action at school contributing to fostering their students learning behavior.

1.2. Implications for future research

Our findings suggest that diagnostic competence concerning learning behavior is a multidimensional process that can be measured and fostered based on the tested three-dimensional model. Implications for future research should aim to foster this competence and its predictors at different stages of a teacher's career. Training programs based on the model should be implemented, and curricula and modules in teacher education on this topic should be developed or modified. A generalization of the model and instrument and planned training programs for different target groups are necessary because only grammar school teachers were investigated in this study. Thus, further studies should incorporate other school types such as primary schools. Furthermore, the operationalization of reflected experience needs to be optimized in further studies. The idea is to manipulate reflected experience in training studies by the application of diaries in which teachers are asked to reflect on their diagnosing experiences. Simultaneously, diaries can promote the transfer of training contents into the classroom, which thereby can measured. In addition, a longitudinal analysis across teachers' careers based on a representative sample could help us to learn more about teachers' competence development. Concerning the case scenario, a validation study with objective classroom data is required.

Conclusion

The construct of teachers' diagnostic competence has been broadened in this study by the results of an empirically tested and validated process model that contains important

teacher actions and knowledge for diagnosing students' learning behavior appropriately and for providing didactic action afterwards. A suitable instrument exists for measuring that competence. Important factors for its development are knowledge about diagnosing learning behavior and a teacher's professional self-concept. On that basis, an amelioration of teacher education and further education concerning this competence is indicated. In particular, training programs for teachers who are already on the job and did not profit from the newly modified second phase of teacher education are needed. In addition to the implementation of training programs, reflected experience can be stimulated by the help of diaries in order to aid teachers' diagnostic competence.

References

Abs, H.J. (2007). Überlegungen zur Modellierung diagnostischer Kompetenz bei Lehrerinnen und Lehrern. [Considerations to modelling diagnostic competence of teachers]. In M. Lüders, & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* [*Research of teacher education. Competence development and programme evaluation*], (S. 63 - 84). Münster: Waxmann.

Arnold, K.-H. (1999). Diagnostische Kompetenz erwerben [Acquiring diagnostic competence]. *Pädagogik, 7-8,* 73-77.

Bakkenes, I., Vermunt, J.D. & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction, 20,* 533-548.

Bates, C., & Nettelbeck, T. (2001). Primary school teachers' judgments of reading achievement. *Educational Psychology, 21,* 177–187.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238 – 246.

Berliner, D. C. (1988). *The development of expertise in pedagogy.* Paper presented at the meeting of the American Association of Colleges for Teacher Education, New Orleans, LA.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35,* 463–482.

Brante, G. (2009). Multitasking and synchronous work: complexities in teacher work. *Teaching and Teacher Education, 25 (3),* 430–436

Browne, M.W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21,* 230 – 258.

Bruder, S., Klug, J., Hertel, S., Kelava, A. & Schmitz, B. (submitted). Counselling Parents in Learning Strategies – Modelling an Important Aspect of Teachers´ Professional Competences. *British Journal of Educational Psychology.*

Butler, D. L. & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*, 245-281.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18,* 947-967.

Coladarci, T. (1986). Accuracy of Teacher Judgments of Student Responses to Standardized Test Items. *Journal of Educational Psychology, 78 (2),* 141-146.

Demaray, M. K., & Elliot, S. N. (1998). Teachers' judgments of students' academic functioning: A comparison of actual and predicted performance. *School Psychology Quarterly, 13,* 8 –24.

Edelenbos, P. & Kubenek-German, A. (2004). Teacher assessment: the concept of `diagnostic competence`. *Language Testing, 21,* 3, 259-283.

Epstein, R. M. & Hundert, E. M. (2002). Defining and assessing professional competence. *American Medical Association, 287 (2),* 226-237.

Fend, H. (2006). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen [New theory of school. Introduction to the understanding of educational systems]*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Feinberg, A. B. & Shapiro, E. S. (2003). Accuracy of teacher judgements in predicting oral reading fluency. *School Psychology Quarterly, 18,* 52-65.

Fiedler, K., Walther, E., Freytag, P. & Plessner, H. (2002). *Organizational Behavior and Human Decision Processes, 88 (1),* 527–561.

Ramsey, M.C. & Reynolds, C.R. (2000). Development of a Scientific Test: A Practical Guide. In G. Goldstein & M. Hersen, *Handbook of psychological assessment, 3rd ed.*, p. 21-42. Oxford: Elsevier Science.

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77 (1),* 81–112.

Hecht, S.A. & Greenfield, D.B. (2002). Explaining the Predictive Accuracy of Teacher Judgments of Their Students' Reading Achievement: The Role of Gender, Classroom Behavior, and Emergent Literacy Skills in a Longitudinal Sample of Children Exposed to Poverty. *Reading and Writing: An Interdisciplinary Journal, 15 (7-8),* 789-809.

Helmke, A. & Schrader, F.W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching & Teacher Education, 3 (2),* 91-98.

Helmke, A., Hosenfeld, I., & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften [Comparative tests as an instrument for the improvement of diagnosic competence of teachers]. In R. Arnold, & C. Griese (Hrsg.), *Schulmanagement und Schulentwicklung [Schoolmanagement and school development]* (S. 119-144). Hohengehren: Schneider.

Horstkemper, M. (2004). Diagnosekompetenz als Teil pädagogischer Professionalität [Diagnostic competence as a part of educational professionalism]. In *Neue Sammlung* 2/2004, 201-214.

Hoyle, R. H. & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.*) Structural equation modeling: Concepts, issues, and applications* (pp. 158-176). Thousand Oaks, CA: Sage Publications.

Hu, L.T., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424 – 453.

Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1 – 55.

Jäger, R.S. (2007). *Beobachten, bewerten, fördern. Lehrbuch für die Aus-, Fort- und Weiterbildung [Observe, value, foster. Textbook for education and further education]*. Landau: Empirische Pädagogik.

Klug, J., Bruder, S., Keller, S. & Schmitz, B. (in press). Wie hängen Diagnostische Kompetenz und Beratungskompetenz zusammen? - Eine korrelative Untersuchung [Which is the relationship between diagnostic competence and counseling competence? - A correlate study]. In J. Fleischer & D. Leutner. *Themenheft: Modellierung von Kompetenzen im Bereich der Bildung: Die psychologische Perspektive [Special Issue: Competence Modelling in the area of education: The psychological perspective]*, *Psychologische Rundschau, 2.* Göttingen: Hogrefe.

Koeppen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current Issues in Competence Modeling and Assessment. *Journal of Psychology, 216 (2),* 61-73.

Kuha J. (2004). AIC and BIC: comparisons of assumptions and performance. *Sociological Methods & Research, 33,* 188-229.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review (28),* 49–57.

Kunter, M., Klusmann, U., Dubberke, T., Baumert, J., Blum, W., Brunner, M., Jordan, A., Krauss, S., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2007). Linking aspects of teacher competence to their instruction: Results from the COACTIV project. In M. Prenzel (Ed.), *Studies on the educational quality of schools: The final report on the DFG Priority Programme* (pp. 39-59). Münster: Waxmann.

Kretschmann, R. (2009). *Von der Diagnose zum Förderplan – Diagnostik in pädagogischen Handlungsfeldern [From diagnosing to the planning of fostering - Diagnostics in educational spheres of activity]*. Vortrag anlässlich des Symposiums Beobachten-Erkennen-Verstehen-Begleiten, Graz: Pädagogische Hochschule. http://home.arcor.de/rudolf.kretschmann/Aufs/Aufs%20Uebersichten/Paedagogische%20Diagnostik.htm (letzter Zugriff: 03.03.2009)

Lee, K., Chiu, S.N., Hasselt, van, C.A. & Tong, M. (2009). The Accuracy of Parent and Teacher Reports in Assessing the Vocabulary Knowledge of Chinese Children with Hearing Impairment. *Language, Speech, and Hearing Services in Schools, 40,* 31-45.

Lüdtke, O., Köller, O., Marsh, H.W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology, 30,* 263–285.

Maclellan, E. (2004). Initial knowledge states about assessment: novice teachers' conceptualisations. *Teaching and Teacher Education, 20,* 523-535.

Marcos, J. J. M., Miguel, E. S. & Tillema, H. (2009). Teacher reflection on rction: what is said (in research) and what is done (in teaching). *Reflective Practice, 10 (2),* 191-204.

Marsh, H.W. & Hau, K.-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology, 32,* 151-170.

McCoach, D. B., & Black, A. C. (2008). Evaluation of model fit and adequacy. In Ann. A. O'Connell and D. Betsy McCoach (eds.), *Multilevel modeling of educational data.* (pp. 245-273). Greenwich: Information age Publishing Inc.

McElvany, N., Schroeder, S., Richter, T., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H. & Ullrich, M. (in press). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern [Diagnostic skills of teachers concerning the evaluation of

pupil's achievements and task difficulty using learning media with instructional images]. *Zeitschrift für Pädagogische Psychologie.*

McGinn, T., Jervis, R., Wisnivesky, J., Keitz, S. & Wyer, P. C. (2008). Tips for Teachers of Evidence-based Medicine: Clinical Prediction Rules (CPRs) and Estimating Pretest Probability. *Journal of General Internal Medicine, 23(8),* 1261–1268.

Meisinger, E.B., Bradley, B.A., Schwanenflugel, P.J. & Kuhn, M.R. (2010). Teachers' Perceptions of Word Callers and Related Literacy Concepts. *School Psychology Review, 39 (1),* 54 – 68.

Muthén, L. K. & Muthén, B. O. (2011). *Chi-Square Difference Testing Using the Satorra-Bentler Scaled Chi-Square.* Retrieved May 16, 2011, from http://statmodel.com/chidiff.shtml

Muthén, L. K. & Muthén, B. O. (2010). *Mplus user's guide.* Los Angeles, LA: Muthén & Muthén.

Opdenakker, M. C., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and Teacher Education, 22(1),* 1–21.

Ophuysen, S.v. (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahnempfehlung [Comparison of diagnostic decisions of novices and experts at the example of school career recommendation]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 38 (4),* 154-161.

Perels, F., Dignath, C. & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education, XXIV,* 17-32.

Perry, N. E., VandeKamp, K. O., Mercer, L. K., & Nordby, C. J. (2002). Investigating teacher-student interactions that foster self-regulated learning. *Educational Psychologist*, 37(1), 5-15.

Perry, N. E., Hutchinson, L. & Thauberger, C. (2008). Talking about teaching self-regulated learning: Scaffolding student teachers' development and use of practices that promote self-regulated learning. *International Journal of Educational Research, 47,* 97–108.

Pintrich, P.R. & van de Groot E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82 (1),* 33-50.

Pintrich, P.R., D.A.R Smith, T. Garcia, and W. McKeachie (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ).* University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning, Ann Arbor, MI.

Rammstedt, B & John, O.P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit. [Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality.] *Diagnostica, 51 (4),* 195-206.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In Heijmans, R.D.H., Pollock, D.S.G. & Satorra, A. (eds.), *Innovations in multivariate statistical analysis.* A Festschrift for Heinz Neudecker (pp.233-247). London: Kluwer Academic Publishers.

Schmitz, B., & Wiese, B. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31,* 64-96.

Shulman, L. S., & Shulman, J. H. (2004). How and what teachers learn: a shifting perspective. *Journal of Curriculum Studies, 36,* 257-271.

Sowa, P. A. (2009). Understanding our learners and developing reflective practice: Conducting action research with English language learners. *Teaching and teacher education, 25,* 1026-1032.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz [Accuracy of the judgment of pupil's characteristics by teachers and the construct of diagnostic competence]. *Zeitschrift für Pädagogische Psychologie*, 19 (1/2), 85-95.

Strasser, J., & Gruber, H. (2003). Kompetenzerwerb in der Beratung: Eine kritische Analyse des Forschungsstands [Competence acquisition in consultation: A critical analysis of the state of research]. *Psychologie in Erziehung und Unterricht, 50,* 381-399.

Vogt, F. & Rogalla, M. (2009). Developing Adaptive Teaching Competency through coaching. *Teaching and Teacher Education,* 25, 1051–1060.

Wang, M.C. (1973). The accuracy of teacher's predictions on children's learning performance. *The Journal of Educational Research, 66 (10).*

Weinert, F.E. (2001). Concept of Competence: A Conceptual Clarification. In D. Rychen & L. Salganik (Eds.). *Defining and Selecting Key Competencies* (p. 45-65). Seattle: Hogrefe and Huber.

Wilson, E. B. (1952). *An Introduction to Scientific Research.* New York: McGraw-Hill.

Winter, F. (2006). Diagnosen im Dienst des Lernens [Diagnoses in the service of learning]. In G. Becker, M. Horstkemper, E. Risse, L. Stäudel, R. Werning & F. Winter, *Friedrich Jahresheft XXIV*. Seelze: Friedrich Verlag.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-35). San Diego, California: Academic Press.

Zucchini, W. (2000). An introduction to model selection. *Journal of Mathematical Psychology*, *44*, 41-61.

Footnotes

[1]German teacher education consists of two phases. The first one involves studying at the university, whereas the second one is a provisional teaching period that takes place in teacher training colleges in combination with real teaching practice in the classroom.

[2]In Germany, teachers have to write these plans for their students whose achievements are threatened to break down or who do not pass the class.

Manuscript 2

Can Diagnostic Competence of Teachers be fostered by Training and the Use of a Diary? – Evaluation of a Training Program on Diagnosing Students' Learning Behavior for Secondary School Teachers

Julia Klug, Mara Gerich, Stefanie Jauch, and Bernhard Schmitz
Technische Universität Darmstadt, Darmstadt, Germany

Julia Klug
TU Darmstadt, FB03
Institut für Psychologie, AG Pädagogische Psychologie
Alexanderstraße 10
64283 Darmstadt
06151/16-6932
klug@psychologie.tu-darmstadt.de

The preparation of this paper was supported by grants SEHM 1538/5-2 from the German Research Foundation (DFG) in the Priority Program "Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes" (SPP 1293)

Highlights

- A training program on teachers' diagnostic competence is evaluated
- The use of diaries in addition to the training program is tested
- Pre- and posttest measures are combined with time-series data
- The training program enhances teachers' diagnostic competence
- The diary proves to be an adequate instrument to measure transfer

Abstract

Diagnosing is one of teachers' key tasks at school. So far, the accuracy in diagnosing students' academic achievement has often been investigated. However, there's a request to shift the focus to diagnosing learning behavior and a call for further educational programs in that field. The aim of this study was to develop and evaluate a training program that includes standardized diaries based on a three-dimensional process model of teachers' diagnostic competence. Forty-seven grammar school teachers participated. A subgroup worked on 176 diaries. Pre- and posttest measures were combined with time-series data. Results showed that the training program enhanced teachers' diagnostic competence, especially when it came to actions before and during diagnosing. The diary proved to be an accurate instrument for measuring transfer, but it had no additional intervention effect above and beyond the training program. As the demand for diagnoses of learning behavior and the individual fostering of students increases, these concepts prove to be helpful both in teacher education and further education.

Keywords: training program, diagnostic competence, learning behavior, diary, time-series analysis

In doing their daily jobs, teachers are faced with multitasking and highly complex work (Brante, 2009). In addition to having professional knowledge, giving learner-centered instruction, managing their classes, interacting with students, and being motivated role models (Opdenakker & Van Damme, 2006; Kukla-Acevedo, 2009), diagnosing is one of their key tasks as they are challenged to meet diverse learning needs and to adapt their teaching to heterogeneous academic ability as well as to multiple interests and motivations (Vogt & Rogalla, 2009). Looking at prior empirical research from the 1970s until today, students' academic achievement in particular has been investigated. So far, accuracy in teachers' judgments has been operationalized as their ability to accurately judge their students' achievement or task difficulties. Accuracy has been measured by correlating teachers' judgments with the results of standardized tests (e.g., Coladarci, 1986; Feinberg & Shapiro, 2003; Helmke & Schrader, 1987; Lee, Chiu, van Hasselt, & Tong, 2009; Wang, 1973; Spinath, 2005). In this approach to the assessment of teachers' diagnostic competence, the diagnosis of students' achievement has been the focus. However, there's an ongoing request in the theoretical literature to shift the focus to diagnosing learning behavior, which allows for didactic action afterwards, such as fostering students individually and adapting classes to their needs (e.g., Abs, 2007). Klug, Bruder, Kelava, and Schmitz (under revision) developed and empirically tested a model of teachers' diagnostic competence that accounts for learning behavior, thus closing the gap between previous empirical research and recent theoretical demands. It describes the diagnosis of learning behavior as a three-dimensional process, consisting of a preaction, an action, and a postaction phase. Based on that model, teachers' and student teachers' diagnostic competence was measured, and the results showed that teachers on the job lacked knowledge about important content regarding pedagogical diagnostics. In that study, a high need for fostering diagnostic competence was shown (Klug, Bruder, Kelava, & Schmitz, under revision).

There is a growing awareness of the necessity of assisting teachers in their professional development in general (Bakkenes, Vermunt, & Wubbles, 2010). In particular, there is a call for further educational programs to foster new facets of teachers' diagnostic competence (Klieme et al., 2003). Nevertheless, there are few such programs yet. Therefore, the purpose of this study was to develop and evaluate a training program that includes a standardized diary based on the model of teachers' diagnostic competence by Klug, Bruder, Kelava, and Schmitz (under revision) to foster teachers' diagnostic competence concerning learning behavior.

1. Theory of teachers' diagnostic competence

As stated before, there have been many studies investigating teachers' diagnostic competence as accuracy of diagnosis. However, because of the recent requests to shift the focus, we based both our training program and the standardized diary on a model that accounts for the cross-curricular diagnosis of learning behavior and regards diagnosing as a process (Klug, Bruder, Kelava, & Schmitz, under revision). This model will now be described.

1.1. A process model of teachers' diagnostic competence that accounts for diagnosing learning behavior

The model's domain is teachers' diagnoses and it is context-specific in focusing on diagnoses of pupils' learning behavior both at school and while they learn at home. It addresses questions such as how pupils deal with tasks, how they do their homework, how they learn at home, which learning strategies they are able to apply, and how much they self-regulate their learning. The issue of self-regulated learning can be fostered by the teacher if he diagnosed problems in this area (e.g., Perry, Hutchinson, & Thauberger, 2008).

In the model, diagnosing is conceptualized as a process (Jäger, 2007). Following the denotation in models of self-regulation (Zimmerman, 2000; Schmitz & Wiese, 2006), the model consists of three cyclical dimensions in the diagnosis process, which take place in a preaction, an action, and a postaction phase. The factor structure was empirically tested, and the model fit well and demonstrated superiority over other factor solutions (Klug, Bruder, Kelava, & Schmitz, under revision). Figure 1 illustrates the process model of diagnosing learning behavior.

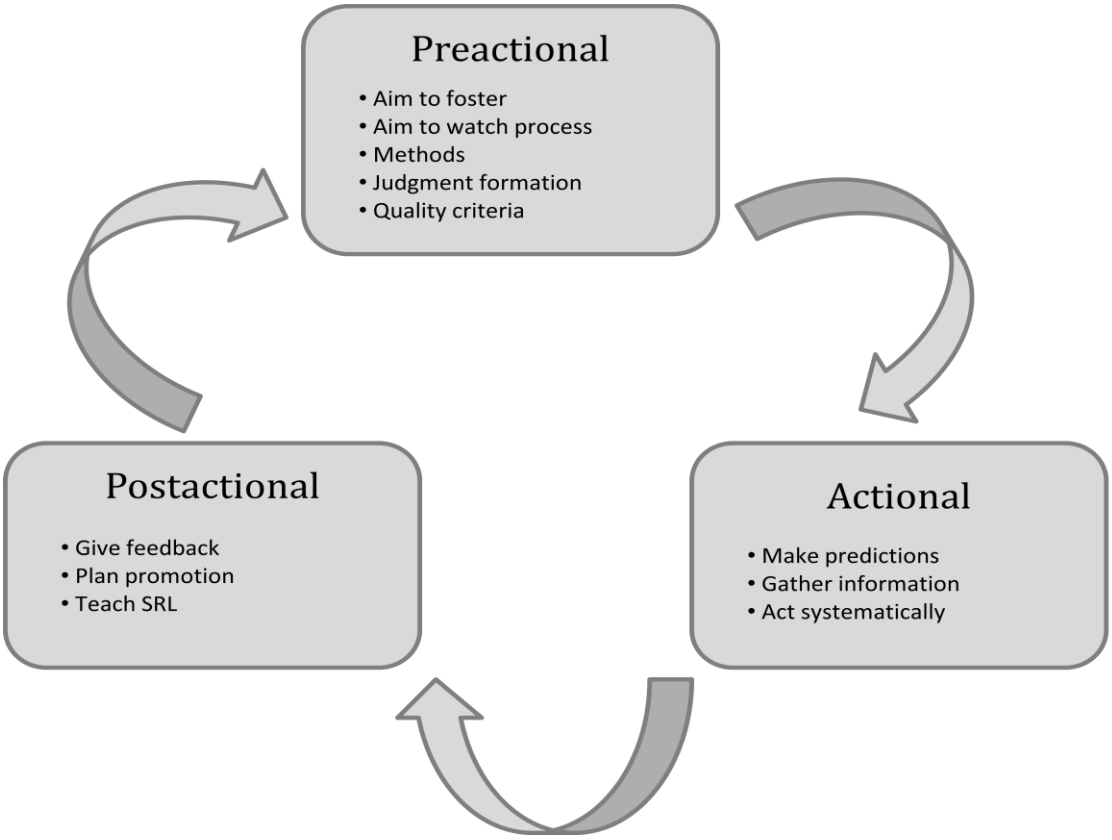


Figure 1. Process model of teachers' diagnostic competence concerning pupils' learning behavior (Klug, Bruder, Kelava, & Schmitz, under revision).

In the following, the three phases of the diagnosis process will be described.

1.1.1. Preaction phase

In the preaction phase, every diagnosing action before summing the information to get an actual diagnosis of a pupil's learning behavior, matters. The teacher needs to set the aim of the diagnosis; for example, by watching one individual student's learning process on a

specific topic and fostering the student based on the diagnosis (Horstkemper, 2004; Abs, 2007). Watching students' learning processes is especially important to have an individual frame of reference and with that to reduce the big-fish-little-pond effect (Lüdtke, Köller, Marsh, & Trautwein, 2005). Furthermore, in the preaction phase, the teacher's basic diagnostic skills are activated (Strasser & Gruber, 2003). These contain knowledge about methods for gathering information (Helmke, Hosenfeld, & Schrader, 2004), quality criteria of tests, and judgment formation (Ophuysen, 2006). The teacher should be familiar with methods to gather information, should know how to deal with these, and should know in which situation which method is the most effective and whether the methods meet quality criteria because teachers need to be able to reason about their assessment practices (Maclellan, 2004). Judgment biases, as a construct from social psychology, can also influence teachers' diagnoses. Fiedler, Walther, Freytag, and Plessner (2002) investigated them in a simulated classroom. Before diagnosing, it is important to be aware of these biases in order to avoid them while diagnosing.

1.1.2. Action phase

In the action phase, the actual diagnostic action takes place. The teacher should act systematically to get a reliable diagnosis—similar to a scientific proceeding when doing quantitative research (Wilson, 1952). The systematic action should begin with making a prediction about a student's development and the possible underlying learning difficulties. This is similar to practicing medicine, where clinicians can use clinical prediction rules (McGinn, Jervis, Wisnivesky, Keitz, & Wyer, 2008). After that, the teacher should gather information from different sources and choose the relevant ones to finally interpret the data and come to a concluding diagnosis. Finally, the teacher can compare real developments with predicted ones in order to possibly change something in the modus operandi for subsequent diagnoses.

1.1.3. Postaction phase

The postaction phase begins right after a diagnosis has been made, when a pedagogical action that follows from the diagnosis should be implemented (Abs, 2007). This includes giving feedback to students and parents. Feedback has been found to significantly influence students' self-regulated learning when given in an effective way (Butler & Winne, 1995; Hattie & Timperley, 2007). Diagnosing also builds the basis for the adequate counseling of parents (Klug, Bruder, Keller, & Schmitz, in press). Additionally, it is important to write plans for the individual student's promotion. These plans should contain the student's skills that need to be fostered, the student's actual skill level, the goals to be reached, and the methods by which these goals are intended to be reached. Finally, adapting a class as a reaction to the diagnosis by means of teaching appropriate learning strategies and self-regulated learning is relevant even for better academic achievements (Pintrich & van de Groot, 1990).

As can be seen in Figure 1, the model has a cyclical character. The three phases are in a timely order and influence each other. Furthermore, a basic assumption is that one diagnosis situation influences consecutive diagnosis situations.

1.1.4. Predictors of diagnostic competence

There are some crucial variables that lead to a high performance in the diagnosis of learning behavior. When the illustrated model was tested, three variables were also tested as predictors of diagnostic competence. Referring to expertise research and various models of teachers' professional development, these variables are: *professional self-concept* consisting of some personal variables, *knowledge* of the topic, and *reflected experience* (e.g., Bakkenes, Vermunt, & Wubbles, 2010; Berliner, 2001; Clarke & Hollingsworth, 2002; Epstein & Hundert, 2002; Kunter et al., 2007; Marcos, Miguel, & Tillema, 2009; Sowa, 2009; Strasser & Gruber, 2003). The three predictors were tested, and two of them, namely professional self-

concept and knowledge, explained a substantial part of the variance of teachers' and teacher students' diagnostic competence (Klug, Bruder, Kelava, & Schmitz, under revision). Reflected experience unexpectedly did not contribute positively to the prediction. Nevertheless, the authors argue that this could be due to the operationalization. Thus, reflection on experienced action can still be seen as crucial for gaining competence. Consequently, not only the model content, but also teaching knowledge, having experiences and reflecting on them, and promoting teachers' professional self-concept should be considered in the training program.

1.2. Training program

In our training program, we incorporated the three phases of the model of teachers' diagnostic competence. Furthermore, we tried to enhance teachers' knowledge, reflected experience, and professional self-concept in diagnosing in order to optimally train their competence. To facilitate reflection on the training content, we additionally implemented a standardized diary. This diary contains questions on every phase of the diagnostic process. According to Webber, Scheuermann, McCall, and Coleman (1993), the continuous registration of one's learning behavior can lead to personal modifications of behavior in a desired direction. Schmitz and Perels (2011) showed that students who work on a learning diary have better results in mathematical problem solving, self-regulation, and self-efficacy than a control group. In this study, the self-monitoring of diagnostic behavior hopefully also leads to a modification of teachers' diagnostic action and self-concept in the desired direction so that they can further enhance their diagnostic competence and better transfer the training content to the classroom.

1.3. Aims

Based on the theoretical assumptions of diagnostic competence concerning learning behavior, we developed a training program and a standardized diary and made the following assumptions: We expected an increase in teachers' diagnostic competence concerning learning behavior if they participated in the training program in contrast to a control group. The increase was expected to be observable in the overall score as well as in the scores of each phase of the diagnostic process and the corresponding variables. Furthermore, we expected an increase in teachers' knowledge, reflected experience, and professional self-concept in diagnostics if they took part in the training program. For the teachers who worked on the diary, we expected an intervention effect in addition to the one from the training program because of the supplementary self-monitoring. In the process data collected by diaries, we expected positive linear trends for each trained variable, reflected experience, and professional self-concept across the training period. Finally, for the diary, we expected a lasting augmentation of scores from a baseline for each training variable to just after the session in which the specific variable was trained.

2. Method

2.1. Participants

Forty-seven secondary school teachers from one German grammar school (Gynamsium), one comprehensive school with a grammar school track, and one teachers' training college in the federal states Hesse and Baden-Württemberg participated. Their mean age was 40.4 years (min = 23, max = 61) and their mean school-teaching experience was 9.89 years (min = 1, max = 38). Thirty-two (68%) were female. Seven had already taken part in a further educational program on diagnostics. For all teachers, participation was voluntary. As

an incentive, teachers were given the opportunity to get credit points and a voucher for a book.

2.2. Design

The longitudinal quasi-experimental design combined pre- and posttest measures with time-series data. There were two experimental groups and one control group. Experimental Group 1 (n = 15) completed the pretest, then got three weekly training sessions, and completed the posttest afterwards. Experimental Group 2 (n = 15) additionally worked on a standardized diary, starting 1 week before the first training session and finishing 1 week after the last session. The control group (n = 17) participated in the pretest and posttest and were offered the opportunity to enroll in a shortened training program afterwards.

2.3. Procedure

In the first session, a pretest containing a test of teachers' diagnostic competence, knowledge, professional self-concept, and reflected experience in diagnostics and some demographic data was given. After that, the training program began. A similar test, supplemented by an evaluation of the training program on reaction levels, was given at the end of the last training session. The training program took place in three weekly 180-min sessions in the schools. To ensure good training conditions, the teachers were trained in three subgroups of no more than 10 participants. Experimental Group 2 additionally worked on a standardized diary in order to self-monitor their diagnostic actions for 4 weeks, starting 1 week before the first training session and ending 1 week after the last one. We expected the self-monitoring to support the training transfer.

2.4. Training program

The training program was developed to foster teachers' diagnostic competence concerning learning behavior and covers the three phases of the diagnostic process. Table 1 summarizes the content of each training session.

Table 1

Content of the Training Sessions

Session	Phase	Content
1	Preaction & action	Pretest Become acquainted with each other Process of diagnosing, approaching systematically Self-assessment Own special case Judgment formation Reflection Homework
2	Preaction & action	Recapitulation Setting aims Making predictions Gathering information Methods Quality criteria Reflection Homework
3	Postaction	Recapitulation Teaching SRL Planning promotion Giving feedback to students and parents Reflection Posttest

Each session resulted in a great deal of activity and reflection. Most importantly, the participants worked on a specific case of one of their own students whom they chose in the first session. Additionally, participants had to do homework for the following session. The content of the previous sessions was recapitulated at the beginning of the consecutive sessions. Furthermore, the trainer made clear that the teachers were the experts at their schools and with their classes, and that with the training program, the teachers would be given methods to choose from, and that they could decide which would best facilitate their diagnosing action.

To begin, the trainers introduced themselves in the first training session and gave an overview of the sessions that would follow. After that, participants worked on the pretest. To motivate participants and to encourage them to become acquainted with each other, an icebreaker game followed. Then the theoretical shift from demanding something from the pupils to assisting the pupils was introduced. With that, the learning objectives were formulated. Teachers' previous knowledge was activated by asking questions about their own diagnosing actions, problems with them, and expectations for the training program using the card and chart technique. Subsequently, the process of diagnosing and the necessity of taking a systematic approach toward diagnosing learning behavior were emphasized with the help of applicable flip charts. On the flip charts, participants were asked to assess their abilities on the training variables with the help of adhesive dots. After every training session, teachers' abilities were assessed with the flip charts in order to monitor the learning progress. A general introduction to diagnostics in the field of education was given before teachers worked on their own special cases on a worksheet for individual work. This case built the foundation to work from in the following sessions. After that, the issue of judgment errors and avoiding them was addressed in three exercises. In group work, the topic was introduced with an exercise in which different perspectives of the same figure were compared and discussed. In a second

exercise, teachers worked in groups on possible judgment errors that were clustered and complemented on a metaplan wall afterwards. Finally, working individually, they reflected on their own possible judgment errors in order to identify and avoid individual problems. At the end of Session 1, as a recapitulation, participants asked and answered questions about the content in a ball game. Subsequently, homework was given, for which the participants had to pay attention to their judgment errors. At the end, participants assessed their abilities again with adhesive dots.

Session 2 began with a short reflection unit in which teachers could talk about experiences and problems in the last week and with their homework. After that, the topics of setting goals, making predictions, and gathering information were introduced. A short theoretical input was followed by a worksheet on which the teachers could work on questions about these topics concerning their own specific cases. Subsequently, methods and quality criteria were dealt with. Participants collected methods in groups, presented them on cards, and discussed their optimal field of application. Methods and their quality criteria were complemented in a presentation by the trainer. In an exercise, participants had to construct their own instrument to gather data in their classes bearing in mind its quality criteria. After that, a theoretical presentation followed; it addressed the issues of checking the formerly made prediction while interpreting the data and applying an individual frame of reference. Subsequently, participants discussed the topic of analyzing discrepancies in order to optimize their diagnostic processes. Finally, the content of the second session was recapitulated and participants again assessed their own abilities on the flip charts. To foster the transfer of the learned content, participants were asked to apply their self-constructed instruments in their classes as homework. Furthermore, they were asked to reflect on their experiences with the application of the instrument, the predictions they made, and how their results could be interpreted, bearing in mind the individual frame of reference of the corresponding student.

Session 3 dealt with the postaction content of the model. Like the second session, it started with a reflection unit. After that, teachers worked in groups on the case of a student with specific learning difficulties. They were instructed to gather intervention possibilities based on their previous knowledge and to present their cases and solutions to the group. Subsequently, the trainer suggested supplementary possibilities for interventions, which were discussed afterwards. The issue of planning promotion began with a discussion because these teachers regularly write plans for promoting at-risk students and therefore have substantial previous knowledge. After a supplementary presentation of the trainer, teachers were instructed to write a plan to promote the student they chose as their own case. Because of the limited time, giving feedback to students and parents was covered in a talk supported by slides with reference to another training program on counseling competence. Finally, the whole diagnostic process was recapitulated in short, and the participants assessed themselves one last time. At the end, they had to work on the posttest before they finally got a brochure that summarized what they covered in the training program.

2.5. Instruments

2.5.1. Measurements of the pretest-posttest evaluation

For the pretest and posttest, a multimethod approach, which had already been established in former studies, was chosen. It consists of three parts: (a) a scenario test with open questions to measure diagnostic competence based on the model, (b) a multiple-choice knowledge test to measure knowledge in diagnostics, and (c) a questionnaire to measure professional self-concept and reflected experience in diagnostics via self-report. The questionnaire also contained some demographic data. The pretest and posttest were conducted within the sessions shortly before the intervention started and after it ended.

2.5.1.1.Case scenario

With a scenario test, we measured diagnostic competence concerning learning behavior based on the model, and that was as close to assessing real action as efficiently as possible. The test consists of a case description of a pupil who has certain difficulties in self-regulated learning, leading to a decline in his achievements. The teacher was asked to imagine being this student's teacher. The case description was followed by 12 open-ended questions, which were formulated with respect to the content of the model. Answers to each question were rated from 0 to 3 points. The raters got a handbook with detailed references about how to rate the answers. For more information on the case scenario, see Klug, Bruder, Kelava, and Schmitz (under revision). In their study, the scenario test proved to be valid for predicting an adequate diagnosis. Inter-rater reliabilities of each question were good with values all between ICC = .67 and ICC = .95.

2.5.1.2.Knowledge test

The knowledge test measures basic knowledge of diagnostics with 11 multiple-choice questions, also constructed with respect to the model content. This test was used in the study by Klug, Bruder, Kelava, and Schmitz (under revision) and proved to be predictive of teachers' diagnostic competence. Item difficulties all fell in an acceptable range between .33 and .82.

2.5.1.3.Questionnaire

The questionnaire surveyed professional self-concept and reflected experience in diagnosing and some demographic data. The scales on professional self-concept and reflected experience were tested for reliability and are further described in the study by Klug, Bruder, Kelava, and Schmitz (under revision).

2.5.1.3.1. Professional self-concept

The scale *professional self-concept* consists of 12 items on a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*) and asks for the teacher's attitude, motivation, and self-efficacy concerning diagnostics and his general conscientiousness. The internal consistency was $\alpha = .77$.

2.5.1.3.2. Reflected Experience

The scale *reflected experience* consists of eight items again on a 6-point Likert scale from 1 (*strongly disagree*) to 6 (*strongly agree*) and asks whether the teacher reflects on the diagnostic experiences he makes. It also had a good internal consistency with $\alpha = .75$.

2.5.2. Measurement of the process evaluation

The process evaluation was based on the standardized diagnosis diaries given to teachers in Experimental Group 2. The diary was newly constructed for this study. Like the scenario and knowledge tests, the diary was constructed with reference to the diagnostic process. The items mainly reflect the content of the model (11 items). Additionally, there are five items on professional self-concept, one item on reflected experience, two introduction items on the teacher's actual condition and the actual day in class, and a final question on the teacher's satisfaction with his or her own diagnostic action on the actual day. The introduction question and the final questions were answered on a scale with five smileys showing different affects. The other items were answered on a 6-point Likert scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Items 1 and 12 both measure a teacher's attitude toward diagnosing for each day on which a diary was worked on. The correlation of these two items should serve as an indicator of the diaries' reliability. The internal consistency of the scale comprised of these two items was high ($\alpha = .84$). Item 8 was inverted to control response sets. At the end of the diary, there was an additional open-ended question regarding which

techniques the participants could apply easily in their classes and how they proceeded in applying them. This question should foster reflection and the transfer of the training content to the school setting. Each diary item was formulated as a state, in contrast to questionnaires that usually survey traits. This means that the items asked for what the teacher did and thought on one particular day. When constructing the diary, we took special care with its length so that it would be easy for teachers to work on it regularly. Therefore, we decided to make it no longer than one sheet of paper, which could be filled out in 3 to 5 minutes. Table 2 gives an overview of the diary items.

Table 2

Overview of the Diary Items

Number	Variable	Item
1	Professional self-concept: Attitude	Today I focused on correctly estimating the learning behavior of my pupils.
2	Professional self-concept: Motivation	Today I was motivated to find out the causes of my pupils' learning difficulties.
3	Professional self-concept: Self-efficacy	Today I had the feeling of being able to estimate the learning behavior of my pupils well.
4	Professional self-concept: Conscientuousness	Today I proceeded carefully in assessing my pupils' learning behavior.
5	Reflected experience	Today after class, I reflected on whether I assessed my pupils' learning behavior correctly.
6	Act systematically	Today I proceeded systematically when assessing my pupils' learning behavior by considering in which phase of the diagnostic process I am.
7	Judgment formation	Today I explicitly cared about special judgment errors so that they do not bias my assessment.
8	Aim to foster (inverted)	Today I assessed my students exclusively to grade them.
9	Aim to watch process	Today, to judge my pupils' learning behavior adequately, I compared their current learning behavior with their earlier learning behavior.
10	Make predictions	Today I compared my prediction of one of my pupils with the learning behavior shown today in order to correct my impression if necessary.
11	Gather information	To find causes of learning difficulties of one of my pupils, I collected information from different sources today.
12	Professional self-concept: Attitude	Today I particularly concentrated on assessing my pupils' learning behavior correctly.
13	Methods	Today, in order to assess my pupils, I used methods apart from the usual examinations, such as observation sheets, pupils' self-assessments, or exchanges with colleagues.
14	Quality criteria	Today I explicitly paid attention to the objectivity and reliability of my assessment methods.
15	Plan promotion	Today I considered how to write a plan to promote one of my pupils.
16	Give feedback	Today I gave feedback to a pupil or one of his parents on his learning behavior in a constructive way.
17	Teach SRL	In addition to normal class, I taught learning strategies today.

3. Results

3.1. Pretest-posttest comparison

There were no significant pretest differences in the dependent measures of the case scenario and the questionnaire, but there was one in the overall score of the knowledge test, $F(2, 44) = 3.31$, $p < .05$, $\eta^2 = .131$. Because of that difference and to maintain similar methods of analysis for all variables, we chose a method that accounted for the pretest difference. The dependent measures of the pretest and posttest were analyzed using a multivariate one-way ANOVA with group as the independent variable and the pretest-posttest differences of the case scenario, knowledge test, and questionnaire measures as dependent variables. Table 3 gives an overview of the significant differences between groups in their development from pre- to posttest for all measures. Means and standard deviations for these measures are shown in Table 4.

Table 3

Overview of Differences between Groups

Independent variable	Dependent variables	*df*	*F*	η^2
Group	Scenario test: overall score	2, 44	12.34***	.361
	Scenario test: preactional	2, 44	5.48**	.199
	Scenario test: actional	2, 44	6.37**	.224
	Scenario test: postactional	2, 44	2.77#	.112
	Knowledge test: overall score	2, 43	12.31***	.364
	Knowledge test: preactional	2, 44	4.80*	.179
	Knowledge test: actional	2, 44	5.80**	.209
	Knowledge test: postactional	2, 44	1.67	.071
	Questionnaire: professional self-concept	2, 44	10.20***	.317
	Questionnaire: reflected experience	2, 44	5.74**	.207

#$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 4

Means and Standard Deviations of the Dependent Measures on the Pretest and Posttest

Group		EG 2 (training program & diary)		EG 1 (training program)		CG	
		M	*SD*	*M*	*SD*	*M*	*SD*
Scenario test: overall score	Pretest	16.87	5.40	16.80	5.07	17.71	3.29
	Posttest	21.87	4.02	21.87	5.18	16.71	4.14
Scenario test: preactional	Pretest	2.00	0.67	1.84	0.57	2.07	0.32
	Posttest	2.43	0.39	2.33	0.52	1.99	0.45
Scenario test: actional	Pretest	1.11	0.47	1.16	0.68	1.24	0.70
	Posttest	1.78	0.37	1.53	0.52	1.12	0.68
Scenario test: postactional	Pretest	1.18	0.58	1.38	0.84	1.22	0.58
	Posttest	1.47	0.63	1.87	0.71	1.14	0.55
Knowledge test: overall score	Pretest	30.45	4.54	24.51	7.24	28.47	7.09
	Posttest	38.17	4.95	35.81	4.02	30.22	6.98
Knowledge test: preactional	Pretest	2.64	0.81	2.18	0.97	2.70	0.83
	Posttest	3.52	0.34	3.27	0.49	2.91	0.73
Knowledge test: actional	Pretest	2.47	1.01	1.67	0.97	2.07	1.10
	Posttest	3.37	0.99	2.98	1.01	2.14	0.92
Knowledge test: postactional	Pretest	3.29	0.50	2.87	0.81	2.92	0.87
	Posttest	3.50	0.48	3.50	0.48	3.09	0.88
Questionnaire: professional self-concept	Pretest	3.84	0.49	3.88	0.69	4.17	0.53
	Posttest	4.26	0.49	4.34	0.65	4.08	0.50
Questionnaire: reflected experience	Pretest	4.27	0.73	4.33	0.78	4.79	0.80
	Posttest	4.69	0.54	4.86	0.61	4.75	0.70

As Table 3 shows, there were significant differences in the pre-post-difference of the groups for nearly every dependent variable with small to medium effect sizes. For just the postaction phase, there were no significant differences on the scenario test and knowledge test measures, but there was a tendency in the expected direction for the scenario test. Contrasts revealed that for all dependent variables, the significant differences were due to differences between the experimental groups and the control group. As expected, both experimental groups had a significantly higher increase than the control group, but the increase for EG 1 did not differ significantly from the increase for EG 2. Figure 2 illustrates the scenario test results for the overall score as well as for the three phases. The figure shows that EG 1 and EG 2 both improved in their diagnostic competence values overall and for each phase from pretest to posttest in a comparable amount, whereas the control group did not. Results for the knowledge test and questionnaire measures were similar. As the questionnaire measures show, there was also an increase in teachers' professional self-concept and reflected experience in both training groups in contrast to the control group.

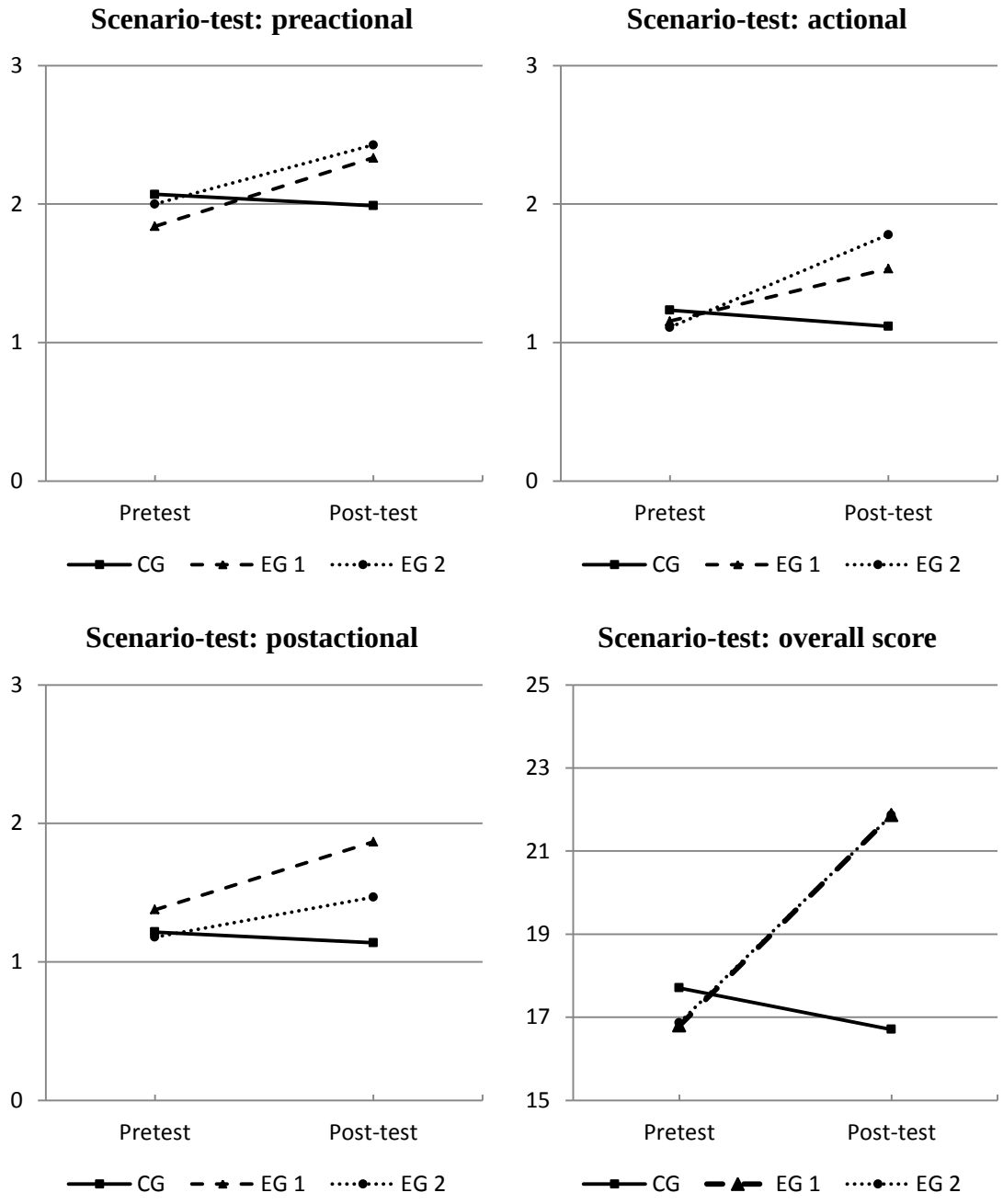


Figure 2. ANOVA groups: scenario-test measures.

3.2. Process evaluation

In addition to the pretest-posttest comparison, the standardized diaries were used to perform time-series analyses. The advantages of studying processes in educational research and especially in the evaluation of training sessions are discussed by Schmitz (2006) and Schmitz and Wiese (2006). We computed trend analyses and interrupted time-series analyses based on the diary data of the EG 2.

One hundred seventy-six of the 300 (59% return rate) distributed diaries were included in the analyses. Trend analyses showed significant linear trends for most of the diary variables. Table 5 gives an overview of the linear trends on the item level for each phase. Figure 3 exemplarily illustrates the significant linear trends for the variables *make predictions* and *plan promotion*. The figure shows that in the course of the training program and the work on the diary, teachers made more and more predictions and cared more and more about promoting students, a trend that continued 1 week after the last training session took place.

Table 5

Linear Trends of Diary Items

Phase/scale	Item	*df*	*F*	*b*
Preaction	Aim to foster	1, 15	7.22*	.04
	Aim to watch process	1, 15	14.63**	.06
	Methods	1, 15	0.24	.01
	Judgment formation	1, 15	16.58**	.11
	Quality criteria	1, 15	22.91***	.10
Action	Make predictions	1, 15	57.44***	.10
	Gather information	1, 15	2.34	.04
	Act systematically	1, 15	13.29**	.09
Postaction	Give feedback	1, 15	0.37	.01
	Plan promotion	1, 15	7.22*	.06
	Teach SRL	1, 15	3.22	-.04
Professional self-concept	Attitude	1, 15	12.78**	.05
	Motivation	1, 15	0.06	.00
	Self-efficacy	1, 15	8.55*	.04
	Conscientiousness	1, 15	5.12*	.06
Reflected experience	Reflected experience	1, 15	7.08*	.06

#$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

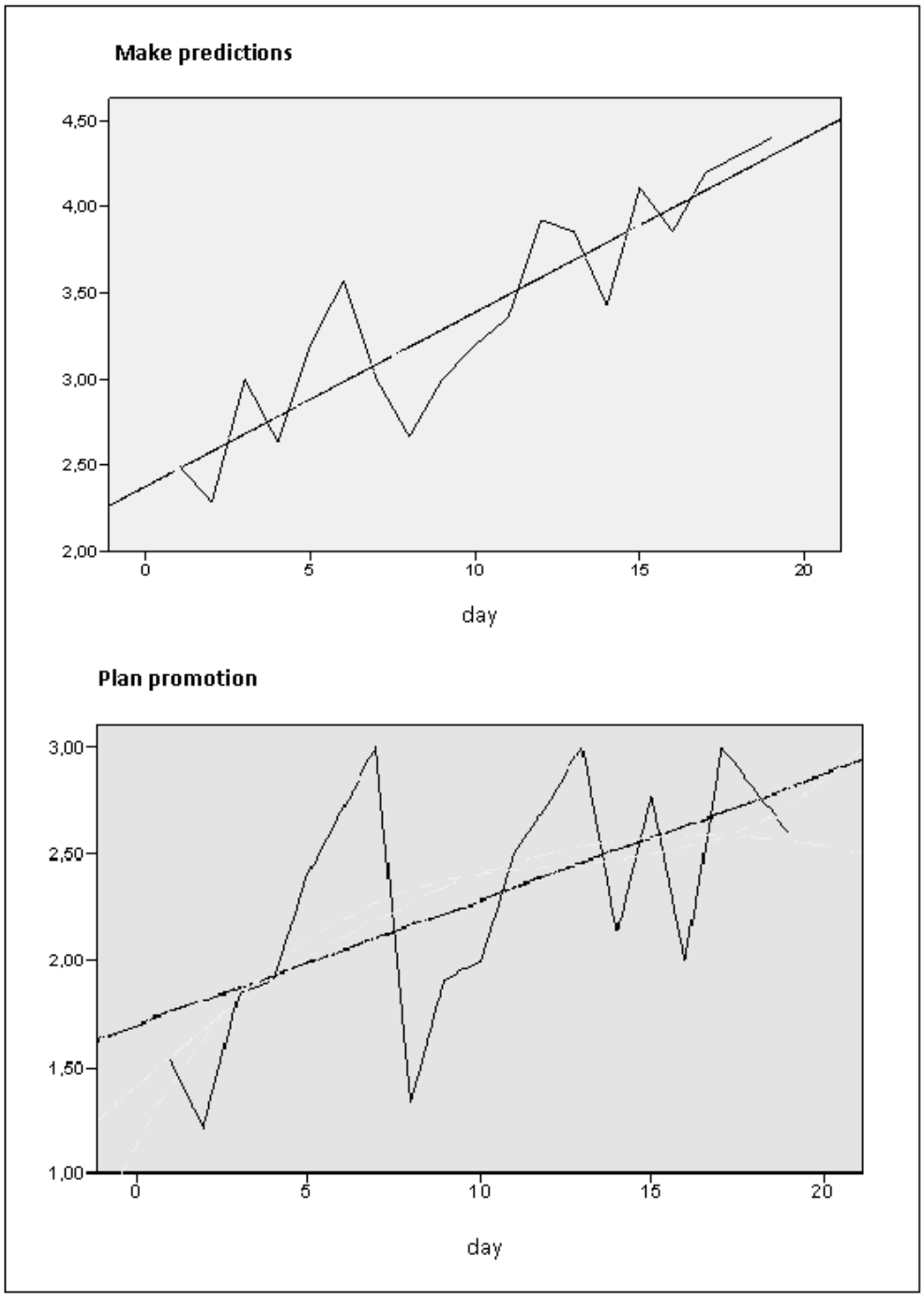


Figure 3. Linear trends for the variables *make predictions* and *plan promotion*.

Interrupted time-series analysis is a well-known procedure for testing intervention effects (Campbell & Stanley, 1963). To conduct this method of analysis, the variables of interest need to be measured for a certain time before the intervention starts (baseline phase).

After this phase, the intervention starts and the measurement points continue (intervention phase). By means of interrupted time-series analysis, the baseline and intervention phases are to be tested for significant differences (Schmitz, Klug, & Schmidt, 2011). In this study, we implemented the baseline by letting the teachers work on the diary 1 week before the training program. With the help of interrupted time-series analyses, we analyzed the effect of each trained variable just after the training session in which the particular content was taught. Furthermore, we looked at the stability of the effects during the following week. Results showed that there was an intervention effect in diary data that remained stable or increased for most of the preaction and action variables of diagnostic competence, but not for postaction variables. Table 6 gives an overview of the intervention effects of each trained variable and assigns the variables to the training session that dealt with them. Figure 4 exemplarily illustrates the intervention effects for the variables *aim to watch process*, which was covered in Session 2 (Day 11), and *judgment formation*, which was covered in Session 1 (Day 6). The figure illustrates that until Day 11, the mean of the baseline phase was significantly lower than the mean after it was trained. Looking at the dashed course, a further augmentation of the values can be recorded for some time after the intervention. The same pattern can be seen for the variable judgment formation, except that this variable had already been trained on Day 6.

Table 6

Intervention Effects of Trained Variables Assigned to the Corresponding Training Session

Phase	Item	Session	*t*	*β*
Preaction	Aim to foster	1	1.30	.31
	Aim to watch process	2	4.28**	.72
	Methods	2	1.49	.35
	Judgment formation	1	5.45***	.81
	Quality criteria	2	3.77**	.69
Action	Make predictions	2	5.57***	.81
	Gather information	2	1.34	.32
	Act systematically	1	3.54**	.66
Postaction	Give feedback	3	-0.79	-.19
	Plan promotion	3	0.91	.22
	Teach SRL	3	-0.47	-.12

*$p < .05$. **$p < .01$. ***$p < .001$.

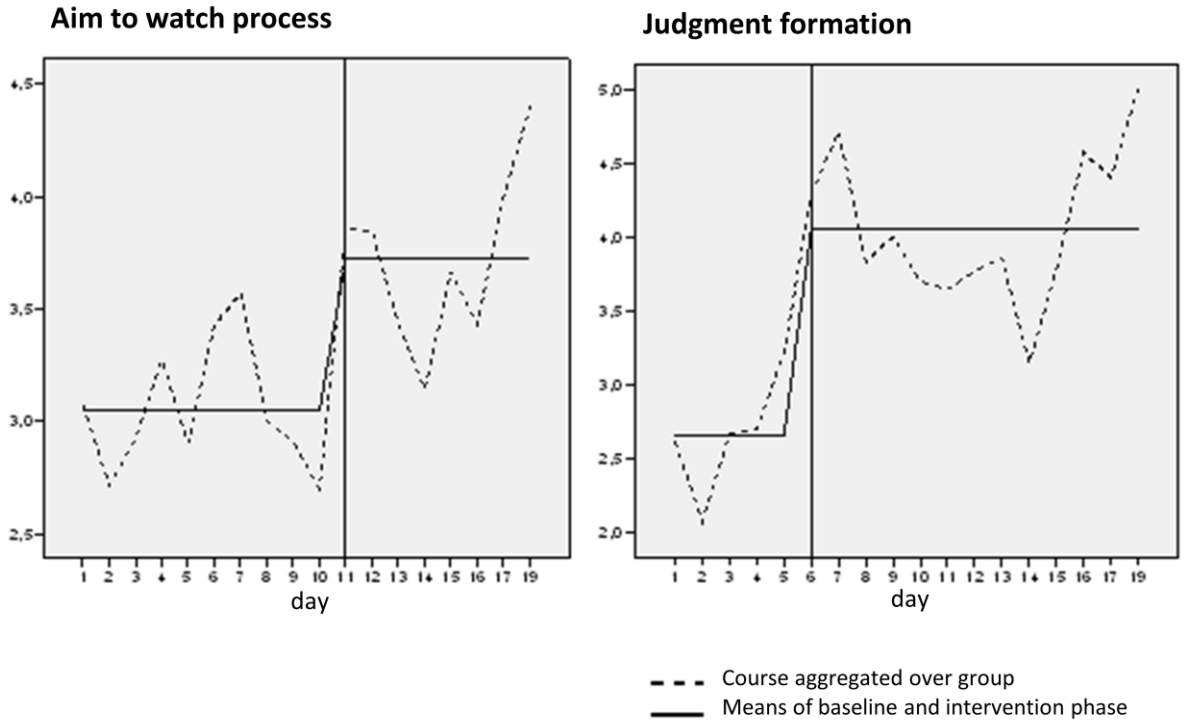


Figure 4. Interrupted time-series analyses for the variables *aim to watch process* (Session 2) and *judgment formation* (Session 1).

4. Discussion

The pretest-posttest comparison showed that, as expected, teachers benefit from the training program in contrast to a control group. Diagnostic competence concerning learning behavior increased in both training groups as reflected by the overall score and the preactional and actional dimensions. Not only did competence increase, but also knowledge in diagnostics showed an increase of the same pattern. In addition, teachers' professional self-concept and reflected experience in diagnostics increased if they took part in the training program, even if those two variables were not explicitly covered. As knowledge and professional self-concept are relevant predictors for the development of diagnostic competence (Klug, Bruder, Kelava, & Schmitz, under revision), it is a great advantage if they increase along the way. Reflection of experiences also is seen as an important condition for the development of competences in expertise research (e.g., Strasser & Gruber, 2003). Thus,

participation in the training program not only increases teachers' diagnostic knowledge and competence directly, but also allows them to develop the basis for further developing their competence. Effect sizes were all small to medium, but in a range that is meaningful, considering the short duration of the training program with only three 180-min sessions.

Unexpectedly, for the postaction measures of diagnostic competence, there was no significant effect due to the training program, but in the pretest-posttest comparison, there was at least a tendency in the expected direction. We suppose that this is due to the short training time in contrast to the complexity of content in the postaction phase. Planning the promotion of students, giving feedback to students and parents, and teaching self-regulated learning covers far too much content for one training session. Thus, we suggest that further studies broaden the postaction content (e.g., in the course of a modularization of the training program with one module consisting of several sessions for each phase). With more time and more practice with the postaction content, we expect the effect to occur.

Concerning the diary, the return rate of 59% can be considered a great success. Teachers accepted the short standardized form of the diary and worked on it constantly over the 4 weeks. The standardization of the items and the length of just one page seem to be about right to motivate teachers to work on it. If it was longer and took more effort, the return rate would probably be lower. As a further source of motivation, teachers who worked on at least 80% of their diaries were given the opportunity to get extra credit points. The use of incentives seems to be a further important way to ensure participants' commitment. Nevertheless, the short and standardized version is good for measurement, analyses, and motivation, but it is at the expense of teachers' reflection and the desired additional intervention effect. The expected additional intervention effect did not occur. A possible explanation may be that there already was much reflection integrated into and stimulated by the training program, so that teachers who did not work on the diary benefitted from self-

monitoring anyway. Maybe it was too much to expect that a short and standardized diary version would offer an intervention effect above and beyond the training. Further studies should investigate whether a longer diary with more open-ended questions to reflect on would lead to the intervention effect. However, the standardized diary proved to be a helpful instrument for measuring the transfer of the trained variables to teachers' everyday work at school. With the diary, teachers reflected on which training content they actually implemented in their classes every day. Furthermore, the process data collected by the diaries gave us much more insight into which diagnostic actions the teachers applied over time and how this augmentation of competence functioned. We expected positive linear trends for the trained variables as well as for teachers' professional self-concept and reflected experience over the 4 weeks. We found these trends for 11 of the 16 measured variables. There was even a significant linear trend for the postaction variable *plan promotion*, which we did not find in the pretest-posttest comparison, even when we analyzed it on the item level. Thus, the diary data provided us with additional information. We also obtained additional information by looking at the interrupted time-series analyses. With it, we were able to illustrate and support an augmentation of the scores from a baseline to just after the session in which the specific variable was trained for many of the trained variables. Furthermore, we were able to see whether the augmentation remained, declined, or even further increased by looking at the course. For the five significant variables, it rose further, supporting a long-term and even delayed effect of the training program and the work on the diary. However, in further studies, the implementation of a follow-up test complementing the pretest-posttest comparison would be great to additionally verify the lasting effects with the other instruments. Postaction variables unexpectedly showed no intervention effect in the process data. In addition to the needed extension of the training program, especially when it comes to postaction content, we assume that the operationalization of the postaction variables in the diary needs to be

improved. Teachers do not and cannot implement these variables (plan promotion, give feedback, teach SRL) every day at school. Thus, the formulation of the items, which, for example, ask whether the teacher taught SRL that day, cannot be answered positively every day even if the intervention was effective. The postaction diary items need to be reformulated perhaps asking for whether the teacher gave thought to those variables that day. Further studies should address whether interrupted time-series analyses will show effects for postaction variables if they are reformulated in this way.

The study design was quasi-experimental, with the advantage of being longitudinal and combining pretest and posttest measures with process data and having a control group. However, the disadvantage is that in the field, no randomized assignment of the teachers to the conditions was possible because the teachers were trained in their schools and had limited time for the appointments. But we did randomly assign which group of participants would be in the experimental and which in the control conditions. Furthermore, in each of the three training groups, half of the participants worked on the diary. In sum, the quasi-experimental design fit the opportunities in the field.

In this study, we chose grammar school and comprehensive school teachers with a grammar school track as the sample. Future research should aim to optimize and evaluate the training program and the diary for other school forms and test whether they can be generalized.

As diagnosing is an everyday task at school and the demand increases to individually foster students, such training concepts should prove to be helpful both in further education and in early teacher education. Thus, the aim should be to increase knowledge in this area by continuing to train teachers and first and foremost, by implementing the concept in teacher education.

References

Abs, H.J. (2007). Überlegungen zur Modellierung diagnostischer Kompetenz bei Lehrerinnen und Lehrern. [Considerations to modelling diagnostic competence of teachers]. In M. Lüders, & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* [*Research of teacher education. Competence development and programme evaluation*], (S. 63 - 84). Münster: Waxmann.

Bakkenes, I., Vermunt, J.D. & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction, 20,* 533-548.

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35,* 463–482.

Brante, G. (2009). Multitasking and synchronous work: complexities in teacher work. *Teaching and Teacher Education, 25 (3),* 430–436

Butler, D. L. & Winne, P.H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65,* 245-281.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18,* 947-967.

Coladarci, T. (1986). Accuracy of Teacher Judgments of Student Responses to Standardized Test Items. *Journal of Educational Psychology, 78 (2),* 141-146.

Epstein, R. M. & Hundert, E. M. (2002). Defining and assessing professional competence. *American Medical Association, 287 (2),* 226-237.

Feinberg, A. B. & Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly, 18,* 52-65.

Fiedler, K., Walther, E., Freytag, P. & Plessner, H. (2002). *Organizational Behavior and Human Decision Processes, 88 (1),* 527–561.

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77 (1),* 81–112.

Helmke, A. & Schrader, F.W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching & Teacher Education, 3 (2),* 91-98.

Helmke, A., Hosenfeld, I., & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften [Comparative tests as an instrument for the improvement of diagnosic competence of teachers]. In R. Arnold, & C. Griese (Hrsg.), *Schulmanagement und Schulentwicklung [Schoolmanagement and school development]* (S. 119-144). Hohengehren: Schneider.

Horstkemper, M. (2004). Diagnosekompetenz als Teil pädagogischer Professionalität [Diagnosis competence as a part of educational professionalism]. *Neue Sammlung* 2/2004, 201-214.

Jäger, R.S. (2007). *Beobachten, bewerten, fördern. Lehrbuch für die Aus-, Fort- und Weiterbildung [Observe, value, foster. Textbook for education and* further education*]*. Landau: Empirische Pädagogik.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. et al. (2003). *Expertise zur Entwicklung nationaler Bildungsstandards. [Expertise for the development of national educational standards]*. Berlin: Bundesministerium für Bildung und Forschung.

Klug, J., Bruder, S., Kelava, A. & Schmitz, B. (under revision). *Diagnostic Competence of Teachers: A Test of a Process Model that Accounts for Diagnosing Learning Behavior.*

Klug, J., Bruder, S., Keller, S. & Schmitz, B. (in press). Wie hängen Diagnostische Kompetenz und Beratungskompetenz von Lehrkräften zusammen? - Eine korrelative Untersuchung. [How do diagnostic competence and counseling competence correlate?]. In J. Fleischer & D. Leutner. *Themenheft: Modellierung von Kompetenzen im Bereich der Bildung: Die psychologische Perspektive [Special Issue: Competence Modelling in the area of education: The psychological perspective], Psychologische Rundschau.* Göttingen: Hogrefe.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review (28),* 49–57.

Kunter, M., Klusmann, U., Dubberke, T., Baumert, J., Blum, W., Brunner, M., Jordan, A., Krauss, S., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2007). Linking aspects of teacher competence to their instruction: Results from the COACTIV project. In M. Prenzel (Ed.), *Studies on the educational quality of schools: The final report on the DFG Priority Programme* (pp. 39-59). Münster: Waxmann.

Lee, K., Chiu, S.N., Hasselt, van, C.A. & Tong, M. (2009). The Accuracy of Parent and Teacher Reports in Assessing the Vocabulary Knowledge of Chinese Children with Hearing Impairment. *Language, Speech, and Hearing Services in Schools, 40,* 31-45.

Lüdtke, O., Köller, O., Marsh, H.W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology, 30,* 263–285.

Maclellan, E. (2004). Initial knowledge states about assessment: novice teachers' conceptualisations. *Teaching and Teacher Education, 20,* 523-535.

Marcos, J. J. M., Miguel, E. S. & Tillema, H. (2009). Teacher reflection on action: what is said (in research) and what is done (in teaching). *Reflective Practice, 10 (2),* 191-204.

McGinn, T., Jervis, R., Wisnivesky, J., Keitz, S. & Wyer, P. C. (2008). Tips for Teachers of Evidence-based Medicine: Clinical Prediction Rules (CPRs) and Estimating Pretest Probability. *Journal of General Internal Medicine, 23(8),* 1261–1268.

Opdenakker, M. C., & Van Damme, J. (2006). Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice. *Teaching and Teacher Education, 22(1),* 1–21.

Ophuysen, S.v. (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahnempfehlung [Comparison of diagnostic decisions of novices and experts at the example of school career recommendation]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 38 (4),* 154-161.

Perry, N. E., Hutchinson, L. & Thauberger, C. (2008). Talking about teaching self-regulated learning: Scaffolding student teachers' development and use of practices that promote self-regulated learning. *International Journal of Educational Research, 47,* 97–108.

Pintrich, P.R. & van de Groot E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82 (1),* 33-50.

Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction, 16,* 433-449.

Schmitz, B., Klug, J. & Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B. Zimmerman & D. Schunck (Eds.): *Handbook of Self-Regulation of Learning and Performance.* New York: Routledge.

Schmitz, B. & Perels, F. (2011). Self-monitoring of self-regulation during math homework behaviour using standardized diaries. *Metacognition and Learning,* DOI: 10.1007/s11409-011-9076-6.

Schmitz, B., & Wiese, B. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31*, 64-96.

Sowa, P. A. (2009). Understanding our learners and developing reflective practice: Conducting action research with English language learners. *Teaching and teacher education, 25,* 1026-1032.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz [Accuracy of the judgment of pupil's characteristics by teachers and the construct of diagnostic competence]. *Zeitschrift für Pädagogische Psychologie,* 19 (1/2), 85-95.

Strasser, J., & Gruber, H. (2003). Kompetenzerwerb in der Beratung: Eine kritische Analyse des Forschungsstands [Competence acquisition in consultation: A critical analysis of the state of research]. *Psychologie in Erziehung und Unterricht, 50,* 381-399.

Vogt, F. & Rogalla, M. (2009). Developing Adaptive Teaching Competency through coaching. *Teaching and Teacher Education,* 25, 1051–1060.

Wang, M.C. (1973). The accuracy of teacher's predictions on children's learning performance. *The Journal of Educational Research, 66 (10).*

Webber, J., Scheuermann, B., McCall, C., & Coleman, M. (1993). Research on Self-Monitoring as Behavior Management Technique in Special Education Classrooms: A Descriptive Review. *Remedial and Social Education, 14,* 38-56.

Wilson, E. B. (1952). *An Introduction to Scientific Research.* New York: McGraw-Hill.

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-35). San Diego, California: Academic Press.

Manuscript 3

Hängen Diagnostische Kompetenz und Beratungskompetenz von Lehrkräften zusammen? Eine korrelative Untersuchung

Julia Klug[1], Simone Bruder[2], Sylvana Keller[1], Bernhard Schmitz[1]

[1] Technische Universität Darmstadt, Darmstadt, Germany

[2] Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt, Germany

In press in *Psychologische Rundschau* (IF 2009: 1.80)

Klug, J., Bruder, S., Keller, S. & Schmitz, B. (in press). Wie hängen Diagnostische Kompetenz und Beratungskompetenz von Lehrkräften zusammen? - Eine korrelative Untersuchung. In J. Fleischer & D. Leutner. *Themenheft: Modellierung von Kompetenzen im Bereich der Bildung: Die psychologische Perspektive, Psychologische Rundschau.* Göttingen: Hogrefe.

Author Note

The preparation of this paper was supported by grants SEHM 1538/5-2 from the German Research Foundation (DFG) in the Priority Program "Competence Models for Assessing Individual Learning Outcomes and Evaluating Educational Processes" (SPP 1293).

Please address correspondence concerning this article to: Julia Klug, TU Darmstadt, FB03, Institut für Psychologie, AG Pädagogische Psychologie, Alexanderstraße 10, 64283 Darmstadt, Germany. Phone: +496151/16-6932. Email: klug@psychologie.tu-darmstadt.de

Zusammenfassung

Diagnostizieren des Lernverhaltens von Schülern und Beraten von Schülern und Eltern sind zentrale Aufgaben des Lehrerberufes. Im Modell der Beratungskompetenz von Lehrkräften von Bruder (2011) bildet Diagnostizieren eine zentralen Kompetenzdimensionen. Genauso spielt Beraten im Prozessmodell der Diagnostischen Kompetenz (Klug, Bruder & Schmitz, 2010) eine zentrale Rolle. Die Diagnose kann in Form eines Beratungsgespräches rückgemeldet werden. Allerdings wurde die Beziehung zwischen beiden Kompetenzen bisher noch nicht empirisch untersucht. An einer Stichprobe von N = 293 Versuchspersonen (n = 93 Lehramtsstudierende, n = 107 Lehrkräfte im Vorbereitungsdienst, n = 93 Lehrkräfte der gymnasialen Mittelstufe), deren Diagnostische Kompetenz und Beratungskompetenz mittels Fallszenarien erhoben wurde, war es nun möglich, diesen Zusammenhang korrelativ zu prüfen. Es konnte ein statistisch signifikanter Zusammenhang der Kompetenzen sowohl auf Ebene der Gesamtscores ($r = .21$, $p < .01$) als auch für wichtige Indikatoren (z. B. Wissen über Diagnostik mit Wissen über Beratung $r = .23$, $p < .01$) nachgewiesen werden. In einer latenten Regression ließ sich die Beratungskompetenz durch die Diagnostische Kompetenz signifikant vorhersagen ($\beta = .38$, $p < .01$). Regressionsanalysen getrennt für die 3 untersuchten Teilgruppen zeigen, dass ein Zusammenhang nur für die Gruppe der Lehrkräfte besteht. Wenngleich durch die korrelativen Analysen kein Kausalzusammenhang nachgewiesen werden kann, ist doch anzunehmen, dass eine gründliche Diagnostik einem guten Beratungsgespräch zeitlich vorausgeht, so dass der postulierte Zusammenhang zwischen Diagnostischer Kompetenz und Beratungskompetenz nun auch empirisch gezeigt werden konnte.

Schlüsselwörter: Diagnostik, Beratung, Kompetenz

Abstract

Diagnosing students' learning behavior and counseling students and parents are central tasks of teachers. In the model of teachers' counseling competence from Bruder (2011), diagnosing learning behavior is an important dimension. Counseling also plays an important role in the process model of teachers' diagnostic competence (Klug et al., 2010). Diagnosing is especially relevant to counseling as it permits a teacher to provide accurate feedback. However, the relationship between diagnostic- and counseling competence has not yet been tested empirically. Within a sample of N = 293 participants (n = 93 grammar school teachers, n = 107 student teachers in the 2nd phase of their education and n = 93 student teachers in the 1st phase of their education) we measured diagnostic- and counseling competence using both case scenarios. Consequently, we were able to test the correlative relationship between both competences. We found a statistically significant correlation between them on the level of total scores ($r = .21$, $p < .01$) as well as on the level of selected predictors of the competences (e.g. knowledge in diagnosing and knowledge in counseling $r = .23$, $p < .01$). In latent regression analysis, we predicted counseling competence significantly based on diagnostic competence ($\beta = .38$, $p < .01$). When we computed the regression separately for the 3 groups, we could only make a significant prediction for the group of teachers. Even though the correlative analyses do not allow for causal interpretations, we assume that a solid diagnosis precedes a good counseling session. Consequently, our data shows empirically that the postulated relationship between diagnostic- and counseling competence exists.

key words: Diagnosing, Counseling, Competence

Im pädagogisch-psychologischen Kontext Schule sind sowohl die Beratungs- als auch die Diagnostische Kompetenz von Lehrkräften in den Standards der Lehrerbildung festgehalten (KMK, 2004) und somit zentrale Aufgabe aller Lehrkräften. Beratungsanfragen in der Schule richten sich nicht nur an Lehrkräfte mit einer spezifischen Ausbildung zum Beratungslehrer, sondern an alle Lehrkräfte. Sie sind Hauptträger der Beratung im deutschen Schulsystem (Landesinstitut für Schule und Weiterbildung, 1998). An Schulen wird zudem ein wachsender Beratungsbedarf konstatiert, wobei insbesondere die Lernberatung an Bedeutung gewinnt (Schnebel, 2007). Auch das Diagnostizieren gehört zu den für das Unterrichten unerlässlichen Kompetenzen (Helmke, Hosenfeld & Schrader, 2004). Vogt und Rogalla (2009) sowie Helmke (2003) beschreiben darüber hinaus, dass gezieltes Diagnostizieren für einen adaptiven, auf die Bedürfnisse der Schüler angepassten Unterricht notwendig ist. Insbesondere das Diagnostizieren des Lernverhaltens der Schüler ist eine wichtige Grundlage um die Schüler effektiv und individuell zu fördern, passende Lernstrategien zu vermitteln und die Eltern adäquat zu beraten. Lehrerkompetenzen als relativ neuer Forschungsgegenstand wurden bisher nur einzeln untersucht. Deren Zusammenspiel kann jedoch auch von Bedeutung sein. Im Modell der Beratungskompetenz von Lehrkräften von Bruder (2011) bildet das Diagnostizieren eine von 4 Kompetenzdimensionen. Genauso spielt das Beraten im Prozessmodell der Diagnostischen Kompetenz von Klug et al. (2010) eine zentrale Rolle in der Phase, nachdem eine Diagnose getroffen wurde. Die getroffene Diagnose sollte dann in Form eines Beratungsgespräches an Eltern und Schüler/innen rückgemeldet werden. Diagnostizieren kann als Grundlage einer guten Beratung angesehen werden. Ein guter Diagnostiker muss nicht zwangsläufig auch ein guter Berater sein, aber es ist ein gemeinsamer Anteil zu erwarten. Dieser wurde allerdings bisher noch nicht empirisch nachgewiesen. Die vorliegende Studie schließt diese Forschungslücke und stellt den Zusammenhang zwischen beiden Lehrerkompetenzen dar.

Theorie

Beratungskompetenz

Neben dem Unterrichten, Erziehen und Beurteilen gehört die Beratung von Schülern und Eltern zu den Kernaufgaben von Lehrkräften an Schulen (KMK, 2004). Beratungswissen wird in aktuellen Modellen zum professionellen Lehrerhandeln explizit als Kompetenzbereich aufgeführt (z. B. Baumert & Kunter, 2006). Auch von Elternseite besteht ein ausgeprägter Wunsch nach Beratung durch Lehrkräfte, da die Eltern sich unsicher fühlen, wie sie ihr Kind beim Lernen unterstützen können (Wild, 2003). 94 % der Lehrkräfte hingegen fühlen sich durch ihre Ausbildung nicht gut auf Beratungsaufgaben im Schulalltag vorbereitet (Hertel, 2009 und fühlen sich in Beratungsgesprächen mit Eltern oft überfordert (Hitzinger, 1987); insbesondere diejenigen Lehrkräfte, die die Fortbildungssituation als unzureichend beurteilen, arbeiten seltener mit Eltern zusammen (Wild, 2003). Andererseits zeigen Studien die Wichtigkeit einer intensiven Zusammenarbeit zwischen Schule und Elternhaus (z.B. Epstein & van Voorhis, 2001). Cox (2005) konnte in einer Meta-Analyse zeigen, dass Interventionen zur Kooperation zwischen Elternhaus und Schule nicht nur wirkungsvoll sind um die Leistung, sondern auch das Verhalten der Kinder in der Schule zu unterstützen.

Trotz der festen Verankerung von Beratung im Schulsystem zeigt sich in der Forschung eine geringe Anzahl empirischer Befunde zu dieser Thematik. Die aktuellen Studien von Hertel (2009) und Aich (2006) betonen die Wichtigkeit der Beratungsarbeit von Lehrkräften und zielen mit Hilfe von Trainingsmaßnahmen auf die Förderung der Gesprächsführungs- und Kommunikationskompetenz ab, die sich signifikant verbessern lässt (Aich, 2006; Hertel, 2009). Was bislang dennoch fehlte, war eine theoretisch und empirisch verankerte Definition der Beratungskompetenz von Lehrkräften (vgl. Hertel, 2009; Strasser & Gruber, 2003). Basierend auf empirisch überprüften Ansätzen und aktueller Literatur (Hertel, 2009; Schwarzer & Buchwald, 2006; Strasser & Gruber, 2003) wurde ein Modell der

Beratungskompetenz entwickelt und empirisch überprüft (Bruder, 2011). Es wird der Domänenspezifität (Klieme & Leutner, 2006) gerecht, bezieht sich ausschließlich auf den Bereich Lernberatung (d. h. Beratung zur Unterstützung von Lernprozessen und selbstreguliertem Lernen) und beinhaltet vier Kompetenzdimensionen (siehe Tabelle 1). Eine ausführliche Beschreibung des Modells und der Skalen findet sich in Bruder, Keller, Klug und Schmitz (2011). Es wurde weiterhin überprüft, welche Faktoren Einfluss auf die Beratungskompetenz haben. Diese sind *Wissen über Beratung und Lernen* sowie die *reflektierte Erfahrung* im Bereich Beratung. Zur Messung der Beratungskompetenz bzgl. des Lernverhaltens wurden ein Szenariotest mit offenen Fragen zum Verhalten im Beratungsgespräch und ein Situational-Judgment Test (SJT) eingesetzt, die noch näher beschrieben werden. Um ein gutes Beratungsgespräch zum Thema Lernverhalten führen zu können, ist es notwendig, das Lernverhalten auch angemessen diagnostizieren zu können.

Tabelle 1

Vier-Dimensionales Modell der Beratungskompetenz

Berater-Skills	**Diagnostizieren**[1]/ Pädagogisches Wissen	Kooperation/ Perspektiv-übernahme	Bewältigung
Aktives Zuhören	Lösungssuche	Kooperatives Handeln	Kritikfähigkeit
Paraphrasieren	Problemdefinition	Perspektiv-übernahme	Umgang mit schwierigen Situationen
Strukturierung	Strategiewissen	Ressourcen-/ Lösungs-orientierung	
	Zielorientierung		

Anmerkungen. [1]Diagnostikanteil im Modell

Diagnostische Kompetenz

Auch die Diagnostische Kompetenz gilt als eine der Kernkompetenzen von Lehrkräften (Weinert, 2001). Meyer (2004) definiert Diagnostische Kompetenz als die Fähigkeit eines Lehrenden, nach festgelegten Kriterien angemessene Urteile über das Lern- und Leistungsverhalten seiner Schüler abzugeben und berücksichtigt dabei explizit Lernverhalten. In der empirischen Forschung wird die Diagnosefähigkeit von Lehrkräften häufig als Fähigkeit verstanden, Schüler zutreffend einzuschätzen (Schrader, 2001). Diagnostische Kompetenz wurde seit 1970 bis heute in einer Vielzahl von Studien untersucht (z. B. Coladarci, 1986; Helmke & Schrader, 1987; Spinath, 2005; McElvany et. al., in press). Dabei stand meist das Konzept der Diagnosegenauigkeit im Vordergrund. Das bisher etablierte Vorgehen zur Messung besteht darin, dass Lehrkräfte Schülerleistungen oder Aufgabenschwierigkeiten beurteilen, die mit tatsächlich erzielten Ergebnissen in Beziehung gesetzt werden (z. B. Helmke et al., 2004; Schrader, 2001; Spinath, 2005). In Untersuchungen zur Rangordnungskomponente – d. h. der Genauigkeit, mit der Schüler hinsichtlich des zu beurteilenden Merkmals in eine Rangreihe gebracht werden – zeigten sich im Durchschnitt mittelhohe Zusammenhänge (r = .64) und zugleich erhebliche Unterschiede zwischen Lehrkräften (Helmke et al., 2004). Spinath und Höfer (2003) untersuchten mit dieser Analysemethode Schülermerkmale wie Intelligenz, Selbstkonzept, Ängstlichkeit und Lernmotivation. Die Werte für die Rangordnungskomponente waren deutlich niedriger (Intelligenz r = .40, Selbstkonzept r = .39, Ängstlichkeit r =.15, Lernmotivation r = .20), was auf die höheren Urteilsanforderungen bei solchen Merkmalen hindeutet. Hosenfeld, Helmke und Schrader (2002) untersuchten außerdem die Genauigkeit der Lehrereinschätzung von Aufmerksamkeit, Verständnis, Interesse und Unterforderung. Die Lehrkräfte unterschätzten diese Schülermerkmale. In theoretischen Beiträgen zur Diagnostischen Kompetenz wird nun ein Paradigmenwechsel hin zur individuellen Diagnostik und Förderung laut, damit sich

Diagnostik nicht als Selbstzweck versteht (vgl. Horstkemper, 2004; KMK, 2004; Kretschmann, 2009). Auch wenn die Beurteilungsgenauigkeit ein wichtiger Bestandteil einer diagnostischen Tätigkeit ist, sollte sie nicht als einzige Repräsentation Diagnostischer Kompetenz bestehen bleiben. Vielmehr ist es wichtig, den Diagnoseprozess, d. h. das Vorgehen beim Diagnostizieren zu beschreiben und zu optimieren, um das Ziel der Förderdiagnostik zu erreichen. Das empirisch überprüfte Modell der Diagnostischen Kompetenz von Klug et al. (2010) fokussiert die Prozesse beim Diagnostizieren des Lernverhaltens der Schüler und wird damit ebenfalls der Forderung der Domänenspezifität gerecht (Klieme & Leutner, 2006). Beim Diagnostizieren des Lernverhaltens geht es darum zu erkennen, welche Lernstrategien der Schüler bereits anwenden kann und wo er sich im Lernprozess befindet. Insbesondere seine Fähigkeit selbstreguliert zu lernen wird diagnostiziert mit dem Ziel, diese zu fördern. Die Diagnostik von Lernverhalten schließt eine Diagnostik von Lernstörungen aus, da diese nicht im Handlungsspielraum einer Lehrkraft liegt. Vielmehr geht es darum Schwierigkeiten und Verbesserungspotential im Lernverhalten der Schüler zu erkennen, so dass anschließend passende Lernstrategien durch die Lehrkraft vermittelt werden können. Das Modell ist zyklisch angelegt und besteht aus drei Dimensionen, die ebendiesen Diagnoseprozess beschreiben. Die dreidimensionale Struktur konnte mittels konfirmatorischer Faktorenanalyse bestätigt werden (Klug et al., 2010). Die Modellfitwerte zeigten eine sehr gute Passung der dreidimensionalen Struktur ($\chi^2 = 47.704$, $df = 36$, $p = .092$, SRMR = .045, RMSEA = .033, CFI = .954) und auch im Vergleich mit einem ein- und zweidimensionalen Modell wies das dreidimensionale bessere Fitwerte auf. Tabelle 2 fasst die Inhalte der drei Dimensionen des Modells zusammen. Zentrale Prädiktoren im Modell sind das Wissen über Diagnostik und persönliche Voraussetzungen, die wir als *professionelles Selbstkonzept* bezeichnen. Als Alternative zur üblichen Messung der Diagnosegenauigkeit und ihrer Komponenten wurde ein neuartiges Instrument entwickelt, um

das Diagnostizieren von Lernverhalten im Sinne des Modells zu messen. Dieses ist ein Szenariotest, ähnlich dem zur Messung der Beratungskompetenz, der das Vorgehen beim Diagnostizieren des Lernverhaltens mit offenen Fragen zu einem fiktiven Fall eines Schülers abfragt (Klug et al., 2010). Auch im Modell der Diagnostischen Kompetenz bildet die Beratung einen wichtigen Aspekt, insbesondere in der postaktionalen Phase, nachdem das Lernverhalten diagnostiziert wurde. Hier sollte eine Rückmeldung an Schüler und Eltern, möglichst in Form eines Beratungsgespräches, gegeben werden. Dennoch muss es nicht bedeuten, dass ein guter Diagnostiker auch ein guter Berater ist. Beide Kompetenzen haben aber einen gemeinsamen Anteil in dem Sinne, dass ein guter Berater auch ein gewissenhafter Diagnostiker sein muss. Diesen gemeinsamen Anteil gilt es nun nachzuweisen.

Tabelle 2. *Dreidimensionales Modell der Diagnostischen Kompetenz*

Präaktional	Aktional	Postaktional
Zielsetzung Förderdiagnostik	Vorhersagen	**Rückmeldung**[1]
Zielsetzung Prozessdiagnostik	Informationssammlung	Förderplanung
Methoden	Systematisches Vorgehen	Lernstrategien
Urteilsbildung		
Gütekriterien		

Anmerkungen. [1]Beratungsanteil im Modell

Hypothesen

Für Beratung im Bereich der Lernunterstützung ist es wichtig, zunächst eine Diagnostik des Lernverhaltens vorzunehmen, um anschließend auch adäquate Strategien vermitteln zu können. Diese inhaltliche Verknüpfung zwischen Beratung und Diagnostik wird in den beschriebenen Modellen vorgenommen und der Zusammenhang soll empirisch

überprüft werden. Diagnostizieren kann dabei als Grundlage, wenn nicht gar als notwendige (aber nicht hinreichende) Bedingung einer guten Beratung angesehen werden (McLeod, 2003). Folgende Hypothesen liegen der dargestellten Studie zugrunde:

(1) Es besteht ein mittlerer korrelativer Zusammenhang zwischen Diagnostischer Kompetenz und Beratungskompetenz, wenn beide mit der gleichen Methode (verhaltensnahe Szenariotests mit offenem Antwortformat) gemessen werden. Von einem mittleren Zusammenhang (bis $r = .70$) wird aufgrund theoretischer Überlegungen ausgegangen. Beide Kompetenzen beinhalten Aspekte der jeweils anderen. Außerdem ist eine angemessene Elternberatung zum Lernverhalten eines Schülers nur möglich, wenn zuvor eine passende Diagnose getroffen wurde, auf deren Basis das Beratungsgespräch stattfindet. Umgekehrt ist es jedoch denkbar, gut diagnostizieren zu können ohne gleichermaßen ein guter Berater zu sein. Der mittlere Zusammenhang sollte sich genauso mittels latenter Regressionsanalysen zeige, in der die Beratungskompetenz durch die Diagnostische Kompetenz vorhergesagt wird, ohne messfehlerbehaftet zu sein.

(2) Es besteht ein kleiner korrelativer Zusammenhang (bis $r = .50$) zwischen Diagnostischer Kompetenz und Beratungskompetenz, wenn beide mit unterschiedlichen Methoden gemessen werden (Szenariotest mit offenem Antwortformat der Diagnostische Kompetenz und Situational-Judgment Test der Beratungskompetenz).

(3) Ebenso bestehen mittlere korrelative Zusammenhänge (bis $r = .70$) zwischen den Prädiktoren der beiden Kompetenzen (Wissen, Professionelles Selbstkonzept, Reflektierte Erfahrung).

In einer weiterführenden Analyse wird post-hoc untersucht, ob sich die Vorhersage für die drei Teilstichproben (Lehrkräfte, Lehrkräfte im Vorbereitungsdienst, Lehramtsstudierende) unterscheidet.

Methode

Stichprobe

An der Untersuchung nahmen insgesamt N = 293 Probanden aus verschiedenen deutschen Bundesländern teil (n = 93 Lehrkräfte der gymnasialen Mittelstufe, n = 107 Lehrkräfte im Vorbereitungsdienst, n = 93 Lehramtsstudierende des Lehramts an Gymnasien). Daten der drei Teilgruppen wurden erhoben, um ein möglichst breites Erfahrungsspektrum abzudecken. 110 Teilnehmer waren männlich (38 %) und 173 weiblich (59 %), 10 gaben ihr Geschlecht nicht an. 175 (60 %) Teilnehmer waren im Alter von 20 bis 29 Jahren, 66 (23 %) im Alter von 30 bis 39 Jahren, 22 (8 %) im Alter von 40 bis 49 Jahren, 25 (9 %) im Alter von 50 bis 69 Jahren und 3 (1 %) waren 60 Jahre oder älter. 2 Personen machten keine Angabe bzgl. ihres Alters.

Instrumente

Zur Messung der Diagnostischen Kompetenz und der Beratungskompetenz wurden verhaltensnahe Szenariotests eingesetzt, die im Folgenden näher beschrieben werden. Diese erfassen Lehrerkompetenzen nicht nur an Selbstberichtsdaten, sondern auch anhand objektiver, verhaltensnaher Messinstrumente und gehen damit der Forderung der Forschung zum Lehrerberuf nach (Kunter & Klusmann, 2010). Zusätzlich kam für die Messung der Beratungskompetenz eine Kurzversion eines Situational- Judgment Tests zum Einsatz. Die Prädiktoren professionelles Selbstkonzept und reflektierte Erfahrung wurden mittels Selbsteinschätzung erfasst. Das Wissen über Diagnostik und Beratung wurde in einem Wissenstest im Multiple-Choice Format gemessen.

Szenariotest Beratung

Der Szenariotest zur Messung der Beratungskompetenz besteht aus der ausführlichen Beschreibung eines Falls mit 12 offenen Fragen die sich den vier Dimensionen des Beratungskompetenzmodells zuordnen lassen und in denen konkretes Beratungshandeln

erfragt wird. Die Instruktion und zwei der offenen Fragen können bei Bruder et al. (2011) nachgelesen werden. Die Intra-class-Korrelationen zur Überprüfung der Interrater-Reliabilität bei der Beurteilung der offenen Antworten lagen in vorhergehenden Studien für alle Dimensionen zwischen .68 und .87. Von einer maximalen Gesamtpunktzahl von 29 wurden hier im Mittel M = 14.57 (SD = 4.33) Punkte erreicht.

Situational-Judgment Test Beratung

Zur Messung der Beratungskompetenz wurde außerdem eine Kurzversion eines Situational-Judgment Test (SJT) eingesetzt. Dieser besteht aus kurzen Fallbeispielen verschiedener Beratungssituationen und jeweils 4 Multiple-Choice Antworten, die mögliche Verhaltensweisen des Lehrers in der beschriebenen Situation repräsentieren. Die aus 13 Items bestehende Originalversion des Tests wurde in einer vorherigen Studie getestet (Keller, Bruder & Schmitz, 2010). Die Items orientieren sich ebenfalls an den Skalen und Dimensionen des Beratungskompetenzmodells von Bruder (2011). Die für die Kurzfassung ausgewählten 6 Items korrelieren signifikant mit dem Gesamtscore des Tests (r = .76, p = .01). Von einer maximalen Gesamtpunktzahl von 24 wurden hier im Mittel M = 16.84 (SD = 3.44) Punkte erreicht (Bruder et al., 2011).

Szenariotest Diagnostik

Bei der Messung der Diagnostischen Kompetenz wurde von den üblichen Konzepten in der bisherigen Forschung abgewichen, da hier nicht die Diagnosegenauigkeit, sondern das Vorgehen beim Diagnostizieren von Lernverhalten auf Basis des dreidimensionalen Prozessmodells gemessen werden sollte. Der Szenariotest wurde in gleicher Weise wie der zur Beratungskompetenz konstruiert. Zu dem Szenario wurden offene Fragen entwickelt, die dem Modell der Diagnostischen Kompetenz von Klug et al. (2010) zugeordnet werden können. Der Einsatz von Methoden zur Informationsgewinnung wird zum Beispiel mit folgender Frage erfasst: *Sie haben sich bereits ein Bild von Marcos Arbeitsverhalten im*

Unterricht gemacht indem Sie Marco bei Gruppenarbeiten und in Stillarbeitsphasen systematisch beobachtet haben. Benötigen Sie weitere Informationen? Wenn ja, welche? Woher bekommen Sie diese? Die Antworten wurden anhand eines ausführlichen Bewertungssystems von unabhängigen Urteilern mit je 0 bis 3 Punkten beurteilt. Die ICCs der Fragen im Szenariotest liegen zwischen .67 und .95. Von einer maximalen Gesamtpunktzahl von 36 wurden hier im Mittel M = 14.28 Punkte (SD = 4.59) erreicht.

Wissenstest

Der Wissenstest besteht aus 20 Fragen im Multiple-Choice Format. 4 Fragen erfassen Beratungswissen, 5 Fragen Wissen über selbstreguliertes Lernen und 11 Fragen die Inhalte des Modells der Diagnostischen Kompetenz. Wenn Mehrfachantworten möglich waren, war dies im Test angegeben. Die Itemschwierigkeiten für die Fragen zur Beratung und zum selbstregulierten Lernen liegen in dieser Studie zwischen .20 und .85, der Mittelwert liegt bei M = 5.57 Punkte (SD = 1.88), bei maximal 9 möglichen Punkten. Die Fragen zum Diagnostizieren weisen Itemschwierigkeiten zwischen .33 und .82 auf. Im Mittel wurden M = 28.42 Punkte (SD = 6.14) erreicht, bei einer maximalen Gesamtpunktzahl von 51.

Fragebogen

Mittels Selbsteinschätzung wurden das professionelle Selbstkonzept und die reflektierte Erfahrung bzgl. Diagnostik und Beratung gemessen. Die Skalen wurden selbst konstruiert, da zu diesen Konstrukten bisher keine Instrumente vorlagen. Der Mittelwert der Skala reflektierte Erfahrung/Beratung beträgt M = 3.28 (SD = .75) und der des professionellen Selbstkonzeptes M = 4.45 (SD = .63). Der Mittelwert der Skala reflektierte Erfahrung/Diagnostik beträgt M = 4.30 (SD = .97) und der des professionellen Selbstkonzepts/Diagnostik M = 4.78 (SD = .51) auf einer Likert-Skala von 1 bis 6.Die internen Konsistenzen der Skalen waren alle zufriedenstellend (reflektierte Erfahrung/Beratung, 5 Items, α = .68; reflektierte Erfahrung/Diagnostik, 4 Items, α = .75;

professionelles Selbstkonzept/Beratung, 17 Items, α = .86; professionelles Selbstkonzept/Diagnostik, 12 Items, α = .77).

Ablauf

Über einen Zeitraum von neun Monaten wurden Gymnasien und Kooperative Gesamtschulen mit gymnasialem Zweig sowie Studienseminare und Universitäten, die Lehramtsstudierende ausbilden, zur Rekrutierung von Teilnehmern kontaktiert. Jeder Teilnehmer bekam als Belohnung einen Buchgutschein im Wert von zehn Euro. Die Lehrkräfte hessischer Schulen hatten die Möglichkeit, Fortbildungspunkte für die Teilnahme an der Erhebung zu bekommen. Es wurde Anonymität und Vertraulichkeit zugesichert. Die Instrumente wurden sowohl in einer Online-Version als auch in einer Paper-Pencil-Version eingesetzt.

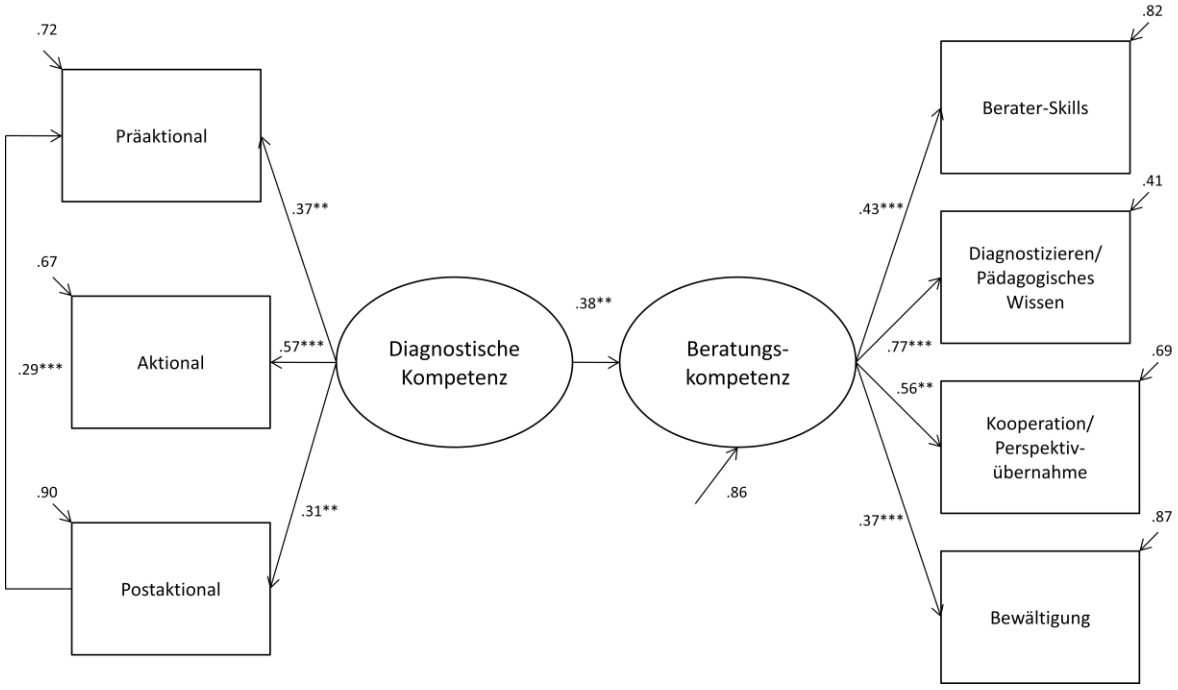
Analysen

Um die Zusammenhänge zwischen den beiden Kompetenzen zu ermitteln, wurden zunächst Korrelationsanalysen gerechnet. In einem nächsten Schritt wurde eine latente Regressionsanalyse berechnet, um die Beratungskompetenz und ihre Dimensionen durch die Diagnostische Kompetenz und deren Dimensionen vorherzusagen. Zusätzlich wurde in einer weiterführenden Analyse überprüft, ob sich die Vorhersage für die drei Teilstichproben unterscheidet. Hierzu wurde eine manifeste Regressionsanalyse mit dem Prädiktor Gesamtwert im Fallszenario/ Diagnostischen Kompetenz und dem Kriterium Gesamtwert im Fallszenario/ Beratungskompetenz mit der Grouping-Variable Gruppe durchgeführt.

Ergebnisse

Der angenommene korrelative Zusammenhang zwischen Diagnostischer Kompetenz und Beratungskompetenz zeigte sich sowohl auf Ebene des Gesamtscores der Szenariotests mit $r = .21$ ($p < .001$) als auch für die Prädiktoren. So korreliert das Wissen über Diagnostik mit dem Wissen über Beratung zu $r = .23$ ($p < .001$). Weiterhin korreliert das professionelle

Selbstkonzept/Diagnostik mit dem professionellen Selbstkonzept/Beratung zu $r = .31$ ($p < .001$) und die reflektierte Erfahrung/Diagnostik mit der reflektierten Erfahrung/Beratung zu $r = .25$ ($p < .001$). Werden die beiden Kompetenzen mit unterschiedlichen Methoden erfasst, korreliert die Diagnostische Kompetenz gemessen mit dem offenen Szenariotest mit der Beratungskompetenz gemessen mit dem SJT immer noch zu $r = .14$ ($p < .05$). Die Berechnung der latenten Regression zur Vorhersage der Beratungskompetenz und ihren Dimensionen durch die Diagnostische Kompetenz und ihre Dimensionen zeigt, dass eine signifikante Vorhersage möglich ist. Die Fitwerte des latenten Regressionsmodells erweisen sich als sehr gut (CFI = 1, RMSEA = .000, SRMR = .027). Der χ^2-Test ist nicht signifikant ($\chi^2 = 8.848$; $df = 12$, $p = .716$). 14 % der Varianz der Beratungskompetenz werden durch die Diagnostische Kompetenz und deren Dimensionen vorhergesagt. Das latente Regressionsmodell ist in Abbildung 1 dargestellt. Es gibt einen gerichteten Pfad zwischen den Indikatoren *postaktional* und *präaktional* der Diagnostischen Kompetenz, der durch den zyklischen Charakter des Diagnoseprozesses theoretisch begründet und im ursprünglichen Modell von Klug et al. (2010) so enthalten ist.



* p < .05; ** p < .01; *** p < .001

Abbildung 1. *Latente Regression zur Vorhersage der Beratungskompetenz durch die Diagnostische Kompetenz.*

Um weiterführend zu analysieren, ob sich die Vorhersage für die drei Gruppen Lehrkräfte, Lehrkräfte im Vorbereitungsdienst und Lehramtsstudierende unterscheidet, wurde eine manifeste Regressionsanalyse mit dem Prädiktor Gesamtwert im Fallszenario zur Diagnostischen Kompetenz und dem Kriterium Gesamtwert im Fallszenario zur Beratungskompetenz mit der Grouping-Variable Gruppe (1 = Lehrkräfte, 2 = Lehrkräfte im Vorbereitungsdienst, 3 = Lehramtsstudierende) durchgeführt. Dabei ergaben sich Unterschiede in der Vorhersage des Gesamtwertes der Beratungskompetenz durch den Gesamtwert der Diagnostischen Kompetenz. Für die Gruppe der Lehrkräfte konnte die Beratungskompetenz hochsignifikant (β= .515, $p < .001$) vorhergesagt werden mit einer Varianzaufklärung von 27 %. Für die Lehrkräfte im Vorbereitungsdienst und die Lehramtsstudierenden hingegen ist eine Vorhersage nicht möglich.

Diskussion

Ziel der dargestellten Studie war es, den bisher nur theoretisch angenommenen Zusammenhang zwischen Diagnostischer Kompetenz und Beratungskompetenz auch empirisch nachzuweisen. Bezogen auf die erste Hypothese zeigte sich nicht wie angenommen ein mittlerer, sondern ein kleiner Zusammenhang, wenn beide mit der gleichen Methode erfasst werden. Die Vorhersage der Beratungskompetenz durch die Diagnostische Kompetenz mittels latenter Regressionsanalyse zeigt hypothesenkonform, dass eine Vorhersage möglich ist. Der in der zweiten Hypothese, bei unterschiedlicher Erhebungsmethode, postulierte kleine erwies sich als sehr kleiner Zusammenhang. Im Sinne eines Multitrait-Multimethod-Ansatzes war ein kleinerer Zusammenhang als bei der Messung mit gleicher Methode anzunehmen. Dass der erwartete Zusammenhang in Bezug auf die erste Fragestellung nur klein und nicht mittel, und der in der zweiten sehr klein statt klein ist, kann dahingehend interpretiert werden, dass es sich bei den beiden Kompetenzen um unterschiedliche Konstrukte mit einem gemeinsamen Anteil handelt. Dieser ist jedoch geringer als zuvor von den Autoren angenommen. Da sich der Zusammenhang trotz unterschiedlicher Erhebungsmethoden signifikant zeigt, kann davon ausgegangen werden, dass der Zusammenhang nicht methodenbedingt zustande gekommen sein kann, sondern tatsächlich besteht. Bezüglich der dritten Hypothese, der angenommenen mittleren Zusammenhänge zwischen den Prädiktoren, zeigen die Ergebnisse auch hier kleine signifikante Zusammenhänge. Es kann geschlussfolgert werden, dass die Voraussetzungen zur Weiterentwicklung in beiden Kompetenzen einander gegenseitig bedingen und die Lehrkräfte, die viel Wissen, Erfahrung und eine positive Einstellung zu der einen Kompetenz aufweisen, diese auch in der jeweils anderen Kompetenz haben. Da die Prädiktoren jedoch jeweils mit gleicher Methodik gemessen wurden, ist ein Methodenfaktor nicht auszuschließen. Es kann jedoch davon

ausgegangen werden, dass sich die kleinen Zusammenhänge in ähnlicher Weise wie auf der Ebene der Kompetenzen auch mit unterschiedlichen Messmethoden noch nachweisen lassen.

Zur Berechnung der weiterführenden getrennten Analyse der drei Gruppen wäre als Analysemethode eine Mehrgruppenanalyse auf latenter Ebene angemessener gewesen. Diese war aufgrund der Stichprobengröße jedoch nicht durchführbar. Bei der stattdessen druchgeführten Regressionsanalyse mit den manifesten Variablen der beiden Gesamtwerte im Szenariotest und der Grouping-Variable Gruppe zeigt sich, dass die Vorhersage nur für die Gruppe der Lehrkräfte gelingt. Das Regressionsgewicht ist hier größer als in der Gesamtgruppe und mit einer Varianzaufklärung von 27 % kann ein nicht unerheblicher Teil der Varianz der Beratungskompetenz der Lehrkräfte durch deren Diagnostische Kompetenz erklärt werden. Für die beiden Gruppen der Lehrkräfte im Vorbereitungsdienst und der Lehramtsstudierenden gelingt hingegen keine signifikante Vorhersage. Die Aufteilung nach Gruppen zeigt, dass der Zusammenhang nicht aufgrund des Ausbildungsstandes der Stichprobe zustande gekommen sein kann. Es stellt sich die Frage, warum der Zusammenhang nur für Lehrkräfte festzustellen ist. Betrachtet man die erreichten Niveaus der drei Gruppen und deren Streuung, so zeigt sich auch im varianzanalytischen Vergleich zwischen den Gruppen, dass sich Lehrkräfte und Lehrkräfte im Vorbereitungsdienst signifikant von Studierenden in dem Sinne unterscheiden, dass sie ein höheres Niveau als Studierende erreichen (Bruder, Klug, Hertel & Schmitz, 2010). Dies ist der Fall beim Gesamtwert der Diagnostischen Kompetenz (M_{Lehrer} = 14.61, SD_{Lehrer} = 4.06, M_{LiV} = 15.54, SD_{LiV} = 4.16, $M_{\text{Stud.}}$ = 12.56, $SD_{\text{Stud.}}$ = 4.93) und Beratungskompetenz (M_{Lehrer} = 17.31, SD_{Lehrer} = 2.94, M_{LiV} = 17.31, SD_{LiV} = 3.50, $M_{\text{Stud.}}$ = 15.81, $SD_{\text{Stud.}}$ = 3.65) sowie den meisten Dimensionen. Lehrkräfte und solche im Vorbereitungsdienst unterscheiden sich hingegen nicht signifikant voneinander in ihrer Ausprägung der beiden Kompetenzen. Die Streuungen der Teilstichproben unterscheiden sich nicht stark voneinander. Niveau- oder

Streuungsunterschiede scheinen nicht als Erklärung für den fehlenden Zusammenhang in der Stichprobe der Lehrkräfte im Vorbereitungsdienst zu genügen. Dennoch ist anzunehmen, dass sich der Zusammenhang erst mit größerer Berufserfahrung zeigt. Im Mittel erreichen die Lehrkräfte im Vorbereitungsdienst zwar das gleiche Niveau in beiden Kompetenzen wie Lehrkräfte, die bereits im Beruf sind, allerdings scheint es keinen systematischen Zusammenhang zwischen der Entwicklung beider Kompetenzen zu geben. Vielmehr scheinen einige Lehrkräfte im Vorbereitungsdienst bereits gute Diagnostiker zu sein, während sie in Beratung noch nicht so gut abschneiden. Gleiches gilt auch umgekehrt. Hingegen zeigt sich, dass diejenigen Lehrkräfte, die gute Diagnostiker sind, überwiegend auch gute Berater sind und umgekehrt. Betrachtet man Einzelfälle und deren Differenz zwischen den erreichten Werten im Fallszenario/ Diagnostik und im Fallszenario/ Beratung, zeigt sich, dass diese größer ist bei Lehrkräften im Vorbereitungsdienst. Wird ein Kriterium von mindestens sechs Punkten Unterschied zwischen den Variablen *Gesamtpunktzahl Fallszenari/Beratung* und *Gesamtpunktzahl Fallszenario/Diagnostik* gewählt, so weisen nur 12.2 % der Lehrer (n = 11) mindestens eine Punktedifferenz von sechs auf, während in der Gruppe der Referendare 35.9 % (n = 37) und in der Gruppe der Studenten 34.4 % (n = 32) mindestens eine Punktedifferenz von sechs Punkten aufweisen. Weiterhin ist beachtenswert, dass sich die Stichprobe der Lehrkräfte im Vorbereitungsdienst aus einer heterogenen Gruppe mit Teilnehmern aus dem ersten bis vierten Hauptsemester zusammensetzt. Vor allem die Teilnehmer im vierten Hauptsemester (n = 8) sind für das erreichte Niveau verantwortlich, das sich im Mittel nicht von dem der Lehrkräfte unterscheidet.

Einschränkend ist anzumerken, dass es sich bei der untersuchten Stichprobe um eine selektive Stichprobe handelt, da die Teilnahme freiwillig war und durch Buchgutscheine belohnt wurde. Es ist zu vermuten, dass Lehrkräfte teilnahmen, die generell eher motiviert sind sich weiterzubilden und durch eine relativ hohe Kompetenz in diesen Bereichen

charakterisiert sind. Dagegen sprechen allerdings die erreichten Werte in den Szenariotests, die mit einem Mittelwert von M = 14.57 (SD = 4.33) bei 29 erreichbaren Punkten zur Beratung und einem Mittelwert von M = 14.28 (SD = 4.59) bei 36 erreichbaren Punkten beim Diagnostizieren doch eher moderat ausfallen. Eine verpflichtende Erhebung für alle Lehrkräfte zufällig ausgewählter Schulen wäre aussagekräftiger und wünschenswert für weitere Studien.

Weiterhin ist anzunehmen, dass Diagnostische Kompetenz eine notwendige, aber nicht hinreichende Bedingung für Beratungskompetenz darstellt. Diese These sollte in der Fragestellung dieser Studie und der hier angewandten Methodik noch nicht überprüft werden, bietet jedoch Potenzial für weiterführende, eingehendere Untersuchungen der Art des Zusammenhangs zwischen Diagnostik und Beratung. Wenngleich durch die korrelativen Analysen kein Kausalzusammenhang nachgewiesen werden kann, ist doch anzunehmen, dass eine gründliche Diagnostik einem guten Beratungsgespräch zeitlich vorausgeht, so dass der zuvor oft implizit postulierte Zusammenhang zwischen Diagnostischer Kompetenz und Beratungskompetenz nun auch empirisch gezeigt werden konnte. Vor diesem Hintergrund wäre es ebenfalls wünschenswert, den Zusammenhang zwischen Diagnostizieren von Lernverhalten und Beratung bereits in der Ausbildung stärker zu thematisieren. Auch durch Weiterbildungen für bereits aktive Lehrkräfte sollte Diagnostizieren von Lernverhalten und Beraten von Eltern und Schüler/innen gemeinsam gezielt trainiert werden.

Literatur

Abs, H. J. (2007). Überlegungen zur Modellierung diagnostischer Kompetenz bei Lehrerinnen und Lehrern. In M. Lüders & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* (S. 63-84). Münster: Waxmann.

Aich, G. (2006). *Kompetente Lehrer. Ein Konzept zur Verbesserung der Konflikt- und Kommunikationsfähigkeit.* Hohengehren: Schneider.

Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9,* 469-520.

Bruder, S. (2011). *Lernberatung in der Schule. Ein zentraler Bereich professionellen Lehrerhandelns.* Veröffentlichte Dissertation. Technische Universität Darmstadt. Zugriff am 15.03.2011. Verfügbar unter http://tuprints.ulb.tu-darmstadt.de/2432/1/Dissertation_Bruder_Lernberatung_070311.pdf.

Bruder, S., Keller, S., Klug, J. & Schmitz, B. (2011). Ein Vergleich situativer Methoden zur Erfassung der Beratungskompetenz von Lehrkräften. *Unterrichtswissenschaft, 39,* 123-137.

Bruder, S., Klug, J., Hertel, S. & Schmitz, B. (2010). Messung, Modellierung und Förderung der Beratungskompetenz und Diagnostischen Kompetenz von Lehrkräften. In K. Beck & O. Zlatkin-Troischanskaia (Hrsg.) *Themenheft, Lehrerbildung auf dem Prüfstand* (S. 173-195). Landau: Empirische Pädagogik.

Coladarci, T. (1986). Accuracy of teacher judgments of student responses to standardized test items. *Journal of Educational Psychology, 78* (2), 141-146.

Cox, D. D. (2005). Evidence based interventions using home school collaboration. *School Psychology Quarterly, 20* (4), 473-497.

Eccles, J. S. & Harold, R. D. (1996). Family involvement in children´s and adolescents´ schooling. In A. Booth & J. F. Dunn (Eds.), *Family-schoollinks: How do they affect educational outcomes*? (pp. 3-34). Mahwah: Lawrence Erlbaum.

Epstein, J. L. & van Voorhis, F. L. (2001). More than minutes: Teachers´ roles in designing homework. *Educational Psychologist, 36*, 181-193.

Helmke, A. (2003). Unterrichtsevaluation: Verfahren und Instrumente. *Schulmanagement, 1* (1), 8-11.

Helmke, A., Hosenfeld, I. & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In R. Arnold & C. Griese (Hrsg.), *Schulmanagement und Schulentwicklung* (S. 119-144). Hohengehren: Schneider.

Helmke, A. & Schrader, F. W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching & Teacher Education, 3* (2), 91-98.

Hertel, S. (2009). *Beratungskompetenz von Lehrern. Kompetenzdiagnostik, Kompetenzförderung und Kompetenzmodellierung*. Münster: Waxmann.

Hitzinger, H. (1987). *Soziale Kompetenz des Lehrers in der Elternarbeit. Analyse und Bewertung eines problemorientierten gruppendynamischen Lehrertrainings zur Vermittlung sozialer Kompetenz in der Lehrerfortbildung.* Frankfurt: Haag & Herchen.

Horstkemper, M. (2004). Diagnosekompetenz als Teil pädagogischer Professionalität. *Neue Sammlung, 44*, 201-214.

Hosenfeld, I., Helmke, A. & Schrader, F.-W. (2002). Diagnostische Kompetenz: Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Unterrichtsstudie SALVE. *Zeitschrift für Pädagogik, 45. Beiheft*, 65-82.

Keller, S., Bruder, S. & Schmitz, B. (2010, September). *Lernberatungskompetenz situativ messen – Validierung eines Situationsbeurteilungstest*. Vortrag auf der 72. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), Jena.

Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik, 52*, 876-903.

Klug, J., Bruder, S. & Schmitz, B. (2010, September). *Diagnostische Kompetenz von Lehrkräften – ein Prozessmodell zum Diagnostizieren des Lernverhaltens.* Vortrag auf der 72. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF), Jena.

Kretschmann, R. (2009, Januar). *Von der Diagnose zum Förderplan – Diagnostik in pädagogischen Handlungsfeldern.* Vortrag anlässlich des Symposiums Beobachten-Erkennen-Verstehen-Begleiten, Graz.

Kultusministerkonferenz (KMK). (2004). *Standards für die Lehrerbildung: Bildungswissenschaften.* Beschluss der Kultusministerkonferenz vom 16.12.2004. Zugriff am 15.03.2011. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf.

Kunter, M. & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften – Methodische Herausforderungen. *Unterrichtswissenschaft, 38* (1), 68-86.

Landesinstitut für Schule und Weiterbildung. (1998). *Fachgutachten. Beratung in der Schule und im Schulsystem. Ergebnisse einer Überprüfung und Anregungen zur weiteren Entwicklung.* Bönen: Verlag für Schule und Weiterbildung.

McElvany, N., Schroeder, S., Richter, T., Hachfeld, A., Baumert, J., Schnotz, W., Horz, H. & Ullrich, M. (in Druck). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung von Schülerleistungen und Aufgabenschwierigkeiten bei Lernmedien mit instruktionalen Bildern. *Zeitschrift für Pädagogische Psychologie.*

McLeod, J. (2003). *An introduction to counselling.* Buckingham: Open University Press.

Meyer, H. (2004). *Was ist guter Unterricht?* Berlin: Cornelsen.

Schnebel, S. (2007). *Professionell beraten. Beratungskompetenz in der Schule*. Weinheim: Beltz.

Schrader, F. W. (2001). Diagnostische Kompetenz von Eltern und Lehrern. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 68-71). Weinheim: Beltz.

Schwarzer, C. & Buchwald, P. (2006). Beratung in Familie, Schule und Beruf. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie - Ein Lehrbuch* (S. 575-612). Weinheim: Beltz.

Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. *Zeitschrift für Pädagogische Psychologie, 19* (1-2), 85-95.

Spinath, B. & Höfer, R. (2003, September). *Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer/innen und das Konstrukt der diagnostischen Kompetenz.* Vortrag bei der 9. Fachtagung Pädagogische Psychologie, Bielefeld.

Strasser, J. & Gruber, H. (2003). Kompetenzerwerb in der Beratung: Eine kritische Analyse des Forschungsstands. *Psychologie in Erziehung und Unterricht, 50,* 381-399.

Vogt, F. & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education, 25,* 1051-1060.

Weinert, F. E. (2001). Concept of competence: a conceptual clarification. In D. Rychen & L. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45-65). Seattle: Hogrefe and Huber.

Wild, E. (2003). Lernen lernen. Wege einer Förderung der Bereitschaft und Fähigkeit zu selbstreguliertem Lernen. *Unterrichtswissenschaft, 31*, 2-5.

Curriculum Vitae

Julia Klug

Persönliche Angaben

Geburtsdatum:	26.12.1983
Geburtsort:	Groß-Umstadt
Nationalität:	deutsch
Familienstand:	ledig

Studium und Schule

Feb. 08 – Aug. 08	**Diplomarbeit** (Note: 1,0) Thema: Diagnostik der Beratungskompetenz von Lehrer/innen - Entwicklung und Validierung verschiedener Instrumente zur multimethodalen Erfassung der Beratungskompetenz	*Technische Universität Darmstadt , Arbeitsgruppe Pädagogische Psychologie*
Okt. 03 – Aug. 08	**Studium der Psychologie** (Gesamtnote: „sehr gut“)	*Technische Universität Darmstadt*
Juni 03	**Allgemeine Hochschulreife** (Note: 1,2)	*Max-Planck-Gymnasium, Groß-Umstadt*

Beruf

Seit Okt. 11	**Universitätsassistentin post-doc**	*Universität Wien, Institut für Angewandte Psychologie: Arbeit, Bildung, Wirtschaft*
Aug. 08 – Dez. 11	**Promotion** zum Thema „Modeling and Training a New Concept of Teachers‘ Diagnostic Competence”	*Technische Universität Darmstadt , Arbeitsgruppe Pädagogische Psychologie*
Apr. 09 – Okt. 11	**Wissenschaftliche Mitarbeitern** auf einer Landesstelle (Vollzeit)	*Technische Universität Darmstadt , Arbeitsgruppe Pädagogische Psychologie*
Aug. 08 – Apr. 09	**Projektmitarbeiterin** Promotionsstelle auf BAT 2a/2 im DFG SPP „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293)	*Technische Universität Darmstadt , Arbeitsgruppe Pädagogische Psychologie*

Publications

Klug, J., Bruder, S., Kelava, A. & Schmitz, B. (under 2nd review). Diagnostic Competence of Teachers: A Test of a Process Model that Accounts for Diagnosing Learning Behavior. *Contemporary Educational Psychology.*

Klug, J., Gerich, M., Jauch, S. & Schmitz, B. (under revision). Can Diagnostic Competences of Teachers be fostered by Training and the use of a Diagnosis Diary? – Evaluation of a Training Program on Diagnosing Students' Learning Behavior for Secondary School Teachers. *Learning and Instruction.*

Klug, J., Bruder, S., Keller, S. & Schmitz, B. (in press). Hängen Diagnostische Kompetenz und Beratungskompetenz von Lehrkräften zusammen? - Eine korrelative Untersuchung. In J. Fleischer & D. Leutner. *Themenheft: Modellierung von Kompetenzen im Bereich der Bildung: Die psychologische Perspektive, Psychologische Rundschau.* Göttingen: Hogrefe.

Klug, J., Müller, C. & Schmitz, B. (in press). Ein Tagebuch für Hauptschullehrkräfte zur Unterstützung der Reflexionsprozesse beim Diagnostizieren. *Themenheft zum Tagebuchsymposium in Empirische Pädagogik.*

Klug, J., Ogrin, S., Keller, S., Ihringer, A. & Schmitz, B. (2011). A plea for self-regulated learning as a process: Modeling, measuring and intervening. *Psychological Test and Assessment Modeling, 53 (1),* 51-72.

Bruder, S., Keller, S., Klug, J. & Schmitz, B. (2011). Ein Vergleich situativer Methoden zur Erfassung der Beratungskompetenz von Lehrkräften. *Unterrichtswissenschaft, 39 (2),* 123-137.

Schmitz, B., Klug, J. & Schmidt, M. (2011). Assessing self-regulated learning using diary measures with university students. In B. Zimmerman & D. Schunck (Eds.): *Handbook of Self-Regulation of Learning and Performance.* New York: Routledge.

Schmitz, B., Klug, J. & Hertel, S. (2011). Using Diaries in Developmental Research: Collecting and Analyzing Longitudinal Diary Data. In B. Laursen, T. Little & N. Card (Eds.), *Handbook of Developmental Research Methods.* New York: Guilford Publications.

Bruder, S., Klug, J., Hertel, S. & Schmitz, B. (2010). Modellierung der Beratungskompetenz von Lehrkräften. In E. Klieme, D. Leutner, & M. Kenk, Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. *56. Beiheft der Zeitschrift für Pädagogik* (S. 274-285). Weinheim: Beltz.

Bruder, S., Klug, J., Hertel, S. & Schmitz, B. (2010). Messung, Modellierung und Förderung der Beratungskompetenz und Diagnostischen Kompetenz von Lehrkräften. In K. Beck & O. Zlatkin-Troischanskaia (Hrsg.) *Themenheft, Lehrerbildung auf dem Prüfstand (S.173-195).* Verlag Empirische Pädagogik: Landau.

Hertel, S. & Klug, J. (2010). Quasi-Experimentelle Versuchspläne. In H. Holling & B. Schmitz (Eds.), *Handbuch der Psychologie: Statistik, Methoden und Evaluation* (S.49-62). Göttingen: Hogrefe.

Klug, J., Bruder, S. & Schmitz, B. (2010). Tagebücher in Praxis und Forschung. In J. Haberstroh (ed.): *Demenz psychosozial behandeln* (S. 359-372). Heidelberg: AKA Verlag.

Bruder, S., Klug, J., Trittel, M. & Schmitz, B. (2010). Messung und Förderung der Beratungskompetenz von Lehrerkräften. In B. Schwarz, P. Nenniger & R. Jäger (Hrsg.), *Erziehungswissenschaftliche Forschung – Nachhaltige Bildung.* Beiträge zur 5. DGfE-Sektionstagung "Empirische Bildungsforschung" / AEPF-KBBB, im Frühjahr 2009, Landau: Verlag Empirische Pädagogik.

Conference Papers

14.09.-16.09.2011 Klug, J., Kutscher, M., Hupp, J. & Schmitz, B. Ein Trainingsprogramm zur Förderung der Diagnostischen Kompetenz von Grundschullehrkräften. Vortrag auf der 13. Fachtagung Pädagogische Psychologie der Deutschen Gesellschaft für Psychologie in Erfurt.

30.08.-03.09.2011 Klug, J., Gerich, M., Jauch, S., Bruder, S. & Schmitz, B. Can Diagnostic Competences of Teachers be fostered by Training and the use of Diary? – Evaluation of a Training Program on Diagnosing Students' Learning Behavior for Secondary School Teachers. Paper presented at the EARLI in Exeter, UK.

28.02.-02.03.2011 Klug, J., Gerich, M., Bruder, S. & Schmitz, B. Ein Tagebuch für Lehrkräfte zur Unterstützung der Reflexionsprozesse beim Diagnostizieren. Vortrag auf der 73. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF) in Bamberg.

13.09. – 15.09.2010 Klug, J., Bruder, S. & Schmitz, B. Diagnostische Kompetenz von Lehrkräften – ein Prozessmodell zum Diagnostizieren des Lernverhaltens. Vortrag auf der 72. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF) in Jena.

13.09. – 15.09.2010 Klug, J., Bruder, S. & Schmitz, B. Wie hängen Diagnostische Kompetenz und Beratungskompetenz zusammen? – Eine korrelative Untersuchung Vortrag auf der 72. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF) in Jena.

13.09. – 15.09.2010 Bruder, S., Klug, J. & Schmitz, B. Lernberatung – ein Vergleich von Studierenden, Lehrkräften im Vorbereitungsdienst und Lehrkräften. Vortrag auf der 72. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF) in Jena.

15.03. – 17.03. 2010 Bruder, S., Klug, J., Hertel, S. & Schmitz, B. Messung, Modellierung und Förderung der Beratungskompetenz und Diagnostischen Kompetenz von Lehrkräften. Vortrag auf dem 22. Kongress der Deutschen Gesellschaft für Erziehungswissenschaften (DGfE) in Mainz.

07.09.-09.09.2009 Bruder, S., Klug, J., Hertel, S. & Schmitz, B. Messung und Modellierung der Beratungskompetenz von Lehrkräften. Vortrag auf der 12. Fachtagung Pädagogische Psychologie der Deutschen Gesellschaft für Psychologie in Saarbrücken.

25.08.-29.08.2009 Bruder, S., Hertel, S., Klug, J. & Schmitz, B. Modelling and Predicting Teachers´Counselling Competence. Paper presented at the EARLI in Amsterdam, The Netherlands.

23.03.-25.03.2009 Bruder, S., Klug, J., Trittel, M. & Schmitz, B. Messung und Förderung der Beratungskompetenz von Lehrkräften. Vortrag auf der 72. Tagung der AG für Empirische Pädagogische Forschung (AEPF) in Landau.

Poster

19.07. – 22.07.2010 Trittel, M., Klug, J. & Schmitz, B. Diagnosing learning behaviour. Towards a Synergy of students´ and teachers´ perspectives. Poster auf der JURE in Frankfurt.

15.03. – 17.03. 2010 Klug, J., Bruder, S., Trittel, M. & Schmitz, B. Modellierung der Diagnostischen Kompetenz von Lehrkräften. Prädiktoren der Kompetenzentwicklung. Poster auf dem 22. Kongress der Deutschen Gesellschaft für Erziehungswissenschaften (DGfE) in Mainz.

15.03. – 17.03. 2010 Trittel, M., Klug, J. & Schmitz, B. (2010). Diagnostik von Lernverhalten aus den Perspektiven von Lehrkräften und von Schülerinnen. Poster auf dem 22. Kongress der Deutschen Gesellschaft für Erziehungswissenschaften (DGfE) in Mainz.

02.10. –03.10.2009 Bruder, S. Klug, J., Schmitz, B. & Bruder, R. Counselling Competence of Teachers: Models and Predictors. Poster auf dem Jahreskolloqium des DFGSchwerpunktprogramms "Kompetenzmodelle" in Göttingen.

7.09.-09.09.2009 Klug, J., Bruder, S. & Schmitz, B. Erste Schritte zur Modellierung der Diagnostischen Kompetenz von Lehrkräften. Poster präsentiert auf der 12. Fachtagung Pädagogische Psychologie der Deutschen Gesellschaft für Psychologie in Saarbrücken.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich gemäß § 9, Abs. 1 der Promotionsordnung der Technischen Universität Darmstadt vom 12. Januar 1990 (in der Fassung der VII. Änderung vom 28. September 2010) die Dissertationsschrift zum Thema „MODELING AND TRAINING A NEW CONCEPT OF TEACHERS' DIAGNOSTIC COMPETENCE" nach bestem Wissen und Gewissen selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ort, Datum	Dipl.-Psych. Julia Klug